DISCUSSION PAPER SERIES

DP16592
(v. 3)

THE RISE OF SCIENTIFIC RESEARCH IN CORPORATE AMERICA

Ashish Arora, Sharon Belenzon, Konstantin Kosenko,
Jungkyu Suh and YIshay Yafeh

ECONOMIC HISTORY AND INDUSTRIAL ORGANIZATION

CEPR

ISSN 0265-8003

THE RISE OF SCIENTIFIC RESEARCH IN CORPORATE AMERICA

Ashish Arora, Sharon Belenzon, Konstantin Kosenko, Jungkyu Suh and YIshay Yafeh

Discussion Paper DP16592
First Published 03 May 2022
This Revision 29 November 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

- Economic History
- Industrial Organization

THE RISE OF SCIENTIFIC RESEARCH IN CORPORATE AMERICA

Abstract

Corporate science in America emerged in the interwar period, as some companies set up state-of-the-art corporate laboratories, hired trained scientists, and embarked upon basic research of the kind we would associate today with academic institutions. Using a newly assembled dataset on U.S. companies between 1900 and 1940 combining information on corporate ownership, organization, research and innovation, we attempt to explain the rise of corporate research. We argue that it was driven by companies trying to take advantage of opportunities for innovation made possible by scientific advances, while facing an underdeveloped academic research system in the United States: In line with this conjecture, we find that firms close to the technological frontier, operating in sectors where American universities were relatively underdeveloped, were more likely to initiate scientific research in response to rising scientific opportunities. We also find that firms that invested in scientific research were large, diversified and operated in concentrated industries, i.e., were able to capture significant rents from the provision of this public good. Indeed, corporate research was positively correlated with novel and valuable patents, and with high market-to-book ratios.

JEL Classification: L2, N12, O3, O31, O32

Keywords: Innovation, Corporate science, Institutional voids

Ashish Arora - ashish.arora@duke.edu
Duke University, Fuqua School of Business and NBER

Sharon Belenzon - sharon.belenzon@duke.edu
Duke University, Fuqua School of Business and NBER

Konstantin Kosenko - kkosta78@gmail.com
The Bank of Israel

Jungkyu Suh - js13220@stern.nyu.edu
New York University, New York University Stern School of Business

YIshay Yafeh - msyafeh@mscc.huji.ac.il
Hebrew University of Jerusalem and CEPR

Acknowledgements
We thank seminar participants at Bar Ilan University, Ben Gurion University, Boston University, the Hebrew University, the University of Pennsylvania, the ASSA 2021 annual meeting, the Strategy Science 2021 Conference, the 2021 Wharton Technology

and Innovation Conference, the World Economic History 2022 Conference, the Academy of Management 2022 annual meeting, the 2022 meeting of the Operations Research Society in Israel as well as Naomi Lamoreaux, Tim Simcoe, Bhaven Sampat and Tarun Khanna for helpful comments and suggestions. We thank Petra Moser and John Graham for providing data on American Men of Science and financial data on CRSP firms respectively, and Gioia Blayer for excellent research assistance. Arora, Belenzon and Suh acknowledge support from the Fuqua School of Business, Duke University. Belenzon and Yafeh gratefully acknowledge financial support from the Israel Science Foundation, Grant No. 963-2020. Kosenko is grateful for support from the Bank of Israel. Yafeh acknowledges support from the Krueger Center at the Hebrew University of Jerusalem School of Business. The views expressed here are the authors' and do not necessarily reflect those of the Bank of Israel. All remaining errors are our own.

The Rise of Scientific Research in Corporate America

Ashish Arora[*] Sharon Belenzon[†] Konstantin Kosenko[‡] Jungkyu Suh[§]
Yishay Yafeh[¶]

In the interwar period, some American firms began to invest in basic scientific research. Using newly assembled firm-level data from the 1920s and 1930s, we find that companies invested in research because inventions increasingly relied on science, but American universities lagged behind both Europe and the scientific frontier. Firms close to the frontier, relying on disciplines which were underdeveloped in American academia, were likely to invest in research, especially if they were large and operated in concentrated industries (could internalize the benefits). Corporate science seems to have paid off, resulting in novel patents and high market valuations for those engaged in research.

---

[*] Duke University, Fuqua School of Business and NBER (ashish.arora@duke.edu)

[†] Duke University, Fuqua School of Business and NBER (sharon.belenzon@duke.edu)

[‡] Bank of Israel (konstantin.kosenko@boi.org.il)

[§] New York University, Stern School of Business (js13220@stern.nyu.edu)

[¶] Hebrew University of Jerusalem, Jerusalem School of Business Administration, CEPR and ECGI (yishay.yafeh@huji.ac.il)

[‖] We thank seminar participants at Bar Ilan University, Ben Gurion University, Boston University, the Hebrew University, the University of Pennsylvania, the ASSA 2021 annual meeting, the Strategy Science 2021 Conference, the 2021 Wharton Technology and Innovation Conference, the World Economic History 2022 Conference, the Academy of Management 2022 annual meeting, the 2022 meeting of the Operations Research Society in Israel as well as Naomi Lamoreaux, Tim Simcoe, Bhaven Sampat and Tarun Khanna for helpful comments and suggestions. We thank Petra Moser and John Graham for providing data on American Men of Science and financial data on CRSP firms respectively, and Gioia Blayer for excellent research assistance. Arora, Belenzon and Suh acknowledge support from the Fuqua School of Business, Duke University. Belenzon and Yafeh gratefully acknowledge financial support from the Israel Science Foundation, Grant No. 963-2020. Kosenko is grateful for support from the Bank of Israel. Yafeh acknowledges support from the Krueger Center at the Hebrew University of Jerusalem School of Business. The views expressed here are the authors' and do not necessarily reflect those of the Bank of Israel. All remaining errors are our own.

1 Introduction

The systematic application of scientific knowledge is arguably the key source of technological advancements, and in turn, of growth in modern economies. Not only are businesses responsible for applying it, but they also generate a substantial portion of scientific knowledge. Businesses invested nearly $90 billion in basic and applied research in the United States in 2020, accounting for nearly half the total domestic investment in research, and nearly three quarters of total domestic R&D (Boroush, 2022; Wolfe, 2022) Yet, at the beginning of the twentieth century, American firms performed very little R&D, and no internal scientific research. By the 1920s the picture was very different (Mowery, 2009). Not only did firms begin to invest considerable amounts in R&D, but some corporations also started engaging in basic scientific research, of the type we would primarily associate with academia today. This paper explores the reasons why and shows that for large firms that were close to the technological frontier and operated in concentrated industries, investment in internal research offered a way to gain competitive advantage: competitors would not be able to readily acquire the needed scientific knowledge from universities.

By the end of World War I, advances in technology were increasingly science-based, especially in industries such as dyes, medicines, plastics, synthetic fibers, oil refining, electricity, and communications. Inventions required more than just trial and error and tinkering; a deeper scientific understanding of the materials and machines in use, and of the natural forces at work, was needed. In contrast to the image of American academia today, American universities at the time were not in a position to provide the required scientific knowledge as they lagged behind the technological frontier in key fields. Firms had to invest in scientific research themselves if they wanted to solve their pressing technological challenges. General Electric (GE), for instance, had exhausted trial and error methods to reduce the blackening that occurred on the surface of the light bulb. Irving Langmuir, an American chemist hired by GE after completing his PhD at the University of Göttingen in Germany, made fundamental discoveries in surface chemistry in the course of diagnosing the source of the blackening as evaporation from the tungsten filament under extremely high temperatures. This led to a radically different solution: instead of trying to create a better vacuum, Langmuir proposed to fill the bulb with inert gases that would scatter the evaporated particles.[1] Elucidating the science behind existing products also allowed firms to develop valuable new products. DuPont, invested in polymer science to improve its products – paints, rubbers and rayon – all of which were based on polymeric materials. This investment yielded blockbuster products such as nylon and orlon.

As we show below, the firms that invested in science, making up for the scientific gap between the U.S. and Europe in fields where the U.S. was lagging, were typically large corporations (in both absolute and market share terms). These firms were large enough to reap substantial benefits from their investments (that is, internalize the provision of science, a classic public good). Because they often operated in concentrated industries, spillover of knowledge to rivals was apparently not a major concern. Consistent with this reasoning, firms investing in scientific research were more likely to produce valuable and breakthrough inventions, and their invest-

---

[1] Langmuir recalled that his work on surface chemistry allowed him to "conclude with certainty that the life of the lamp would not be appreciably improved even if we could produce a perfect vacuum" (Reich, 1983).

ments were also associated with high stock market valuations, especially in fields where U.S. academic research was weak.

The relationship between the state of academic research and corporate investment in science is complex. Internally generated research by companies could complement or substitute for academic science. Moreover, because academic science is potentially available to all firms in a market, the nature of the strategic interactions among competitors matters as well. We develop a simple conceptual framework to study the private returns to investment in research, conditional on the state of public science. We distinguish between scientific knowledge and innovation. Innovation — the introduction of new products and processes — is the source of profits. Scientific knowledge, either from universities or from internal research, reduces the cost of innovation. Leading firms that are more dependent on innovation derive greater returns from investing in research. However, their incentives also depend upon the nature of strategic interactions, as well as on the state of academic science. We show that, under some conditions, the incentives for leaders to invest in internal research may be higher when the supply of academic science is low.[2]

We study these issues empirically using a newly assembled, comprehensive historical dataset covering most of the non-financial sectors in America at the time. Our dataset combines hitherto unavailable information on corporate research and innovation with financial information for U.S. firms in the interwar period. We start with the 200 largest industrial corporations identified by Berle and Means (1932) (hereafter B&M), accounting for about 60% of non-financial corporate assets in

---

[2] In the present paper we ignore the effect of the production of human capital. Arora, Belenzon, and Patacconi (2019) analyze how the joint production of knowledge and human capital conditions the incentive of a single incumbent in a model where the incumbent may potentially buy inventions from startups.

1930. Given the prevalence of control chains during this period, we add the subsidiaries of the B&M firms from Kandel, Kosenko, Morck, and Yafeh (2019). This dataset includes both publicly listed and large private firms. We expand this sample to include all publicly traded firms from CRSP (Graham, Leary, & Roberts, 2015). Importantly, since R&D data are not available for our sample period, we link our firms to U.S. patents, ending with 466 firms that were engaged in inventive activity during the interwar period.

For each firm in the sample, we assemble three measures of investment in scientific research — scientific journal publications, the operation of research labs, and the employment of prominent scientists — and use several patent-based measures as proxies for innovation. We collect data on scientific publications authored by researchers employed at corporations using Microsoft Academic Graph (MAG). We match firms to the *Industrial Research Laboratories of the United States* survey (hereafter the "IRL" directory) of corporate laboratories. We also link prominent scientists listed in the *American Men of Science* series (hereafter the "AMS") to firms. While a substantial number of firms invest in science, their distribution is skewed: a total of 312 firms (67%) in our sample operate a research lab and 201 (43%) publish, but the top 8% of publishing firms account for 75% of all corporate publications, while the 10 largest corporate laboratories account for half of all lab employment. We use the extent to which a firm's patents cite (rely on) science to proxy for its proximity to the technological frontier, or need for science. As technical breakthroughs become more reliant on scientific knowledge, inventions by scientifically advanced companies tend to rely more on basic science (embodied in academic publications) than inventions from companies behind the frontier.[3]

To measure the public science that firms could draw upon, we develop new measures of relative American backwardness to Europe by scientific field. We calculate, for each scientific field, the number of academic publications by European authors, divided by publications by all (U.S. and European) scientists in the same field. We then relate scientific fields to firms, based on the relevance of each field to the firm's patenting activity. In an alternative measure, we calculate the share of backward citations made to European journals (out of total backward citations) by American journals in each scientific field.

Our sample period, ranging from 1926 to 1940, is well suited to the issue at hand. First and foremost, this is the period when the phenomenon of corporate science emerged in the United States. Second, our sample period ends before the onset of World War II, when the U.S. government became deeply involved in scientific development both directly and through military procurement contracts (Gross & Sampat, 2022; Mowery & Simcoe, 2002). During our sample period, the U.S. government had little influence on corporate research and development activities.

Our main empirical findings are as follows. First, to substantiate the basic premise of the paper — the corporate need for science — we show that innovations, especially breakthrough innovations, became more science-based during our sample period.

Second, we establish that corporate investment in scientific research was driven by firms at the technological frontier: Firms with patents citing scientific articles

---

[3] We focus on in-text citations as front-page NPL (Non-Patent Literature) citations were introduced only in 1947 at the U.S. Patent Office. We also measure technological frontier by whether the firm patents first in a patent class and whether the firm's patent is textually distinct from prior art.

(our proxy for proximity to the technological frontier) publish nearly 38 times more in scientific journals than firms whose patents do not cite scientific articles. The difference is similar using industrial (research) lab employment: firms with patents citing science employ around 350 lab personnel, on average, compared to 23 for firms whose patents do not cite science.

Third, firms investing in science relied on scientific fields in which the U.S. was relatively weak at the time (such as select sub-fields of chemistry and physics).

Fourth, firms investing in science tended to be large in absolute terms (as measured by assets), market leaders in relative terms (with large market shares), and operated in concentrated markets.

Finally, corporate investment in science seems to have paid off: firms that invested in research produced important and valuable patents, and corporate investments in science were associated with high stock market valuation.

Admittedly, these results are empirical associations, as there is no exogenous source of variation in the data. However, they support the explanation that the increasing reliance of technological advances on science, along with the inability of American universities to provide the required scientific knowledge, led firms that had reached the technological frontier to invest in scientific research themselves.

We contribute to, and depart from, the existing literature in several ways. Most importantly, we provide systematic empirical evidence on, and novel explanations for, the rise of corporate research in America as a response to the need for science and the relative backwardness of American academia. The existing literature has focused on other explanations. Lamoreaux and Sokoloff (1999) study the internalization of R&D by U.S. firms around the turn of the twentieth century from the perspective of independent inventors and their gradual conversion into salaried R&D employees in larger firms. Nicholas (2010) attributes this transition partly to the increasing complexity of chemical and electrical technology. Other studies suggest that firms may invest in science to increase absorptive capacity (Cohen & Levinthal, 1989; Rosenberg, 1990), to attract competent inventors (Stern, 2004), or to signal their product quality to buyers (Azoulay, 2002; Hicks, 1995). In addition, company histories (Hounshell & Smith, 1988; Jenkins & Chandler, 1975; Maclaurin & Harman, 1949; Reich, 1985) document the various motives behind the establishment of large industrial R&D laboratories.

The present study is related also to the literature on spillovers (Arora, Belenzon, & Sheer, 2021; Bloom, Schankerman, & Van Reenen, 2013), and the question why many U.S. firms have substantially reduced their investments in scientific research in recent decades (Arora et al., 2019; Mowery, 2009). We describe corporate investment in science as a response to gaps in public science. Our findings may therefore suggest that the dramatic growth of university research after World War II affected the private returns to firms undertaking such research, possibly leading to a decline in corporate science in subsequent decades.

The findings in this study relate to, and provide evidence in support of, the missing institutions (or institutional voids) framework in a context hitherto not documented in the literature. Khanna (2000), and Khanna and Palepu (2000) apply this idea to diversified business groups in emerging markets; the present paper finds that this concept is useful in understanding phenomena related to large corporations in the U.S. economy in the first half of the twentieth century. In this respect our results add a historical perspective to contemporary debates regarding the role of large U.S. corporations in advancing science and promoting innovation. Whereas Autor, Dorn, Katz, Patterson, and Van Reenen (2020) describe large U.S. corporations as highly innovative “superstar firms,” others, such as Gutiérrez and Philippon (2020), view them as primarily inefficient entities shielded from competitive pressures.[4] The present study illustrates how large corporations can play an important role in advancing science and, at the same time, how this role is associated with market power.

Finally, the dataset we construct, combining financial information on U.S. corporations and data on corporate science before the Second World War, is one of the most comprehensive of its kind, complementing the recent literature on this era, which primarily focuses on individual inventors (Babina, Bernstein, & Mezzanotti, 2021) and scientists (Moser, Parsa, & San, 2022). Our dataset should open new research on the potential links between corporate research and development, and government policy.

The paper is organized as follows. Section 2 surveys the historical context and Section 3 sets up the theoretical framework for our analyses. Section 4 describes our data. Section 5 presents our econometric specifications and estimation results. Section 6 concludes by summarizing our findings and their implications.

2 The Rise of Corporate Science in America

The beginning of the twentieth century was marked by leaps in scientific opportunity, and a greater reliance of the chemical and electric industries on these new discoveries. In Figure 1 we find that patents in the two decades preceding the

---

[4] A related literature addresses whether government investment in R&D crowds out private investment (David, Hall, & Toole, 2000; Lichtenberg, 1986).

First World War made hardly any citations to the scientific literature. The following period between 1920 and 1940, on the other hand, shows a marked increase in citations to science. Breakthrough chemical patents during this period, for instance, are up to three times more likely to cite a scientific article than non-breakthrough ones, while breakthrough communication patents (telephone, radio & vacuum tube) are more than seven times more likely to do so than non-breakthrough ones.[5] Firms patenting in fields where innovation became more reliant on science also began to invest in science. Figure 2 shows that, by 1920, the chemical industry employed the most scientists (133), compared to only 18 in 1900. Overall lab employment also increases by almost five times between 1927 and 1940, which accords with prior work (Lamoreaux & Sokoloff, 1999; Mowery & Rosenberg, 1998). Figure 2 also shows that corporate publications in scientific journals grew by 11.5 times for chemical firms and close to 29.2 times for firms patenting in electrical engineering between these two periods.

Company histories indicate that in the early years, corporate labs focused on quality control and solving operational problems rather than fundamental science.[6] As innovation became more science based, companies initially looked, as they had in the past, to external suppliers to fill this need. This motivated the establishment of specialized contract research organizations, such as the Mellon Institute in 1913.[7]

---

[5] Breakthrough patents are patents in the top decile of the "importance" measure from Kelly, Papanikolaou, Seru, and Taddy (2021) (KPST hereafter). The authors define importance as the ratio between 10-year forward similarity to other patents and 5-year backward similarity, net of year fixed effects.

[6] Whereas the earliest corporate researchers such as Charles Dudley at the Pennsylvania Railroad Company focused primarily on testing iron and steel for rails, later cohorts of corporate scientists included renowned scientists such as the two Nobel Laureates, Irving Langmuir of GE (Chemistry, 1932) and Clinton Davisson of AT&T (Physics, 1937).

[7] The institute grew steadily in contract revenues ($300,000 to $800,000) between 1919 and 1929. Over the same period, the number of industrial fellows sponsored by firms grew from 83 to 145.

But contract research worked best for generic, well-specified problems, where contracting problems were less severe, and required that the firm itself possess significant research capabilities (Mowery, 1983). Research managers, such as Willis Whitney at GE Research Labs, Frank Jewett at AT&T Bell Labs, C.E.K. Meese at Kodak, and Charles Stine at DuPont, therefore chose to invest internally. The cases of vacuum tube electronics and wireless (radio) technology below illustrate these points.

2.1 Industrial Response to Early Breakthroughs in Electronics

2.1.1 Discovery of the Electron and Thermionic Emissions

The nineteenth century witnessed a steady stream of investigations by Maxwell and Hertz into electricity and magnetism that culminated in the discovery of the electron in 1897 by J.J. Thomson. Though a key implication of this research was that electronic information could be transmitted wirelessly, reliable transmission and reception of electronic signals did not occur until after follow-on scientific discoveries in the 1920s. For instance, Thomas Edison had discovered the discharge of electric currents from lamp filaments to cathodes as early as 1875, but it was only by 1901 that Owen Richardson proved that the currents were formed by electrons escaping the surface of hot filaments. Termed "thermionic emissions," the phenomenon would form the basis of the radio industry and vacuum tube electronics.

In trying to exploit thermionic emissions, companies discovered they also had to develop some of the basic science. Instruments exploiting this phenomenon were developed after Edison's discovery, with the diode (1904 by John Ambrose Flem-

---

Union Carbide's contract with the institute yielded ethylene glycol, an antifreeze, which became a key product for the firm (Servos, 1994, p.223).

ing) and triode (1906 by Lee De Forest) invented in quick succession. Though these devices were promising prototypes for receiving and amplifying signals (relevant for telecommunications) as well as switching and rectifying currents (relevant for electrical devices), they required substantial improvements. Many of the defects could not be removed without understanding the science underlying the technology, a task that universities seemingly left for industry.[8]

2.1.2 Heterogeneous Responses to Breakthroughs: GE & AT&T vs Westinghouse & Western Union

Only the market-leading firms appear to have been willing to invest in electronics research. By 1900, GE controlled nearly 90% of lamp sales (Wise, 1985), while AT&T controlled around half of the telephone exchange market share in 1907 (Mueller, 1997). Quality improvements or cost savings from applying scientific research could be realized over larger output. Beyond scale, these firms also urgently needed improvements in their core product: rising competition and the expiration of the Edison patents in 1894 spurred GE's research efforts to produce the Coolidge tungsten filament (1910) and Langmuir's gas-filled lamps (1913) which contributed to the longevity of the lamp, but also created new opportunities in vacuum tubes (Birr, 1957).[9]

Unlike GE, Westinghouse neither had the scale nor the urgent need to improve

---

[8] At the time of its publication in 1901, Richardson's theory was still being contested by competing hypotheses which argued that thermionic emissions occurred by the interaction of electrons on the filament with the ambient gas inside the lamp. These theories posited that bulb blackening was due to water vapor in the lamp (Langmuir, 1913; Soddy, 1907). Irving Langmuir's first research project at GE Research Labs was to settle this debate.

[9] GE's incandescent light bulbs, for instance, were facing competition from alternative designs pioneered by German chemists such as Carl Auer Welsbach's Osram, Walther Nernst's glower, and Leo Arons' mercury vapor lamp. The American rights to the Nernst glower were purchased by GE's competitor, Westinghouse, for $ 1 million in 1894.

the incandescent bulb. Westinghouse only had 13% of the lamp market. In addition, the antitrust settlement of 1911 whereby GE lamp patents were licensed to Westinghouse also required that technical information be shared between the two firms. This dampened incentives for in-house lamp (and by extension, vacuum tube) research at Westinghouse (Reich, 1992).[10]

As telephony became more science-based, the need for internal scientific capability, instead of merely relying on independent inventors, became apparent to AT&T management. For instance, AT&T had recently lost a legal battle on the loading coil patent — a critical equipment for long distance calls — against a competing inventor at Columbia University (Lipartito, 1989). Even when it managed to acquire rights to inventions, the firm could not exploit them effectively because the fundamental electronics was poorly understood. For instance, in 1913 AT&T acquired the rights to the De Forest Audion, which could detect and receive radio signals (Reich, 1985). However, the Audion's performance was erratic, with blue haze impeding its functions and De Forest's tantalum filaments prone to breakage. Moreover, De Forest had a poor scientific understanding of his own invention, unaware of the triode's potential as an amplifying device (Hughes, 2004). Replacing the filaments with more reliable oxide-coated cathodes and solving the blue-haze problem required the full-time attention of a scientist (H.D. Arnold) and a team of twenty five researchers (including future Nobel Physics laureate Clinton Davisson) for two years (Hoddeson, 1981). AT&T's improved audions were readily modified as amplifiers on the transcontinental telephone service between New York and San

---

[10] Though Westinghouse established a laboratory in Forest Hills in 1916 and published papers, it was only by the late 1930s with the recruitment of Princeton physicist Edward Condon, that it began to conduct research in nuclear medicine and industrial mass spectroscopy (Lassman, 2003).

Francisco in 1915.[11]

In comparison, competing firms in the telecommunications industry, such as Western Union and Postal Systems, were focused on improving wired telegraphs, which relied on legacy technology that was less science-based. Western Union itself was a firm built on earlier technological innovations such as Ezra Cornell's glass insulated telegraph wires in the 1840s. However, it stayed out of the telephone market as part of an 1879 patent settlement with Bell.[12]

AT&T and GE are examples of market leaders in the communication and electrical industries that reached the technological frontier, faced a gap in the scientific understanding of the technology embodied in their products, and sought to fill this gap by investing in scientific research internally in view of the inadequacies of American university research.

2.1.3 Supply of Public Science in Electronics

Despite the need, American universities were unable or unwilling to invest in research in electronics at the scale required. Much of the scientific foundation in electronics until the 1920s was provided by European scientists at institutions such as the Cavendish laboratory in Britain (Faraday, Maxwell, Thomson, Richardson, and Bragg) and German research universities (Hertz, Siemens, Hittorf, Roentgen, Wehnelt, and Braun) (Maclaurin & Harman, 1949).[13] Until the First World War,

---

[11] An improved understanding of wireless technology by Arnold's group also enabled the opening of a wireless relay on this line in the same year.

[12] AT&T employed 26 prominent scientists in the American Men of Science (AMS) directory of 1921, while Western Union employs none.

[13] The British lead in electromagnetism was established early with Maxwell's equations, but also aided by the Royal Society and imperial projects such as the construction of a global telegraph line during the nineteenth century (Hunt, 2021). Germany's traditional strength in chemistry also allowed for the discovery of new rare-earth substances that were applied as new vacuum tube filaments.

American universities were well behind: American universities had published only nine papers per year in the *Transactions of the American Institute of Electrical Engineers*, between 1920 and 1925, accounting for fewer than 10% of total publications (Terman, 1976).[14] Electrical engineering departments were routinely staffed by instructors with only Bachelor's degrees. Though enrollment in electrical engineering programs rose after the end of the First World War, research in the field was still dominated by firms until the end of the 1920s. No doctorates in electrical engineering were awarded at MIT until 1910. No leading American university produced more than two electrical engineering doctorates a year until the 1930s, whereas MIT alone produced thirteen per year between 1950 and 1954.

The absence of American universities from electronics research can be attributed to a combination of a general reluctance to fund research and the high cost of electronics research. American universities before World War II received very little federal funding for research, and spent little as well. For its 1938 report, the National Resources Planning Board under the National Research Council (NRC) surveyed 1,450 American colleges and universities and found that the top 150 spent an average of $333,333 per university on research (National Research Council, 1938). The University of Chicago ($2,557,803 in 1929-30), and the University of California ($2,350,000 in 1928-29) were the top research spenders. By comparison, DuPont alone spent as much as the two universities put together: DuPont's 1925, 1930 and 1935 budgets were $1.99 million, $5.5 million and $6.6 million, respectively (Hounshell & Smith, 1988, p.612). AT&T's R&D expenditures were even larger – the 1925, 1930 and 1935 budgets were $11.7 million, $23.2 million and $15.4

---

[14] Moreover, seven out of the nine annual publications were concentrated in only five universities.

million, respectively (Maclaurin & Harman, 1949, p.158).

Research in the physical sciences increasingly required expensive equipment. Early electrical engineering problems dealt with the operation of dynamos, whose construction required sophisticated machining. The generation of high vacuum — a pre-requisite to studying thermionic emissions — was itself a formidable technical challenge that required Langmuir to adopt and modify Wolfgang Gaede's molecular pump imported from Germany.[15] The case study of electrical engineering suggests that firms began investing in scientific research due to i) new scientific opportunities; ii) their proximity to the technological frontier, and iii) an inadequate supply of public science. The incentives for investment were greatest for market-leaders that reached the technological frontier.

3 Conceptual Framework

The foregoing account of the rise of corporate research stresses three factors: the imperative to innovate for the leading firms, the role of science in facilitating innovation, and the weakness of American university science. To study more formally how these factors interact, we adapt the framework developed in Arora et al. (2021). Whereas they analyze the impact of spillovers, we focus on the differences across firms in the payoffs from innovation and the effect of public science on research investments.

There are two firms, indexed by 0 and 1. There are three stages. In Stage 3, the firms compete in the product market. Their product market performance depends on the quality of their products and the cost of producing them. We assume

---

[15] For a parallel example from chemistry, only the University of Wisconsin and DuPont operated an ultra-centrifuge, Svedberg's Nobel prize-winning scientific instrument used to separate chemical substances by their molecular weights (Cerveaux, 2013).

that cost and quality depend upon their innovation output, d_0 and d_1, respectively. The payoffs from Stage 3 are $\Pi(d_0, d_1)$ and $\tilde{\Pi}(d_1, d_0)$, where the tilde indicates Firm 1. Firms farther from the frontier can increase profits by imitation, reducing production bottlenecks and increasing scale, possibilities that the leaders have already exhausted. Instead, leaders have to introduce new and improved products and processes—to innovate. Firm 0 is closer to the frontier, so that its marginal return from innovation is higher than that of its rival.

In Stage 2, firms choose their innovation output. The cost of innovation for Firm 0 is $\phi(r_0; u)d_0$, where r_0 represents investments in internal scientific research by the firm and u indexes the stock of (relevant) public science. Innovation typically requires the invention of new products and processes. Internal research reduces the cost of invention by guiding the search for inventions in more promising directions. Public science may also guide such search. Innovations may also be based on inventions acquired from independent inventors, other firms or university researchers. Thus, the cost of innovation also depends on the state of public science. We assume that both internal research and public science reduce the unit cost of innovation, i.e., $\frac{\partial \phi}{\partial r_0} < 0, \frac{\partial \phi}{\partial u} < 0$, and diminishing marginal returns to research, $\frac{\partial^2 \phi}{\partial r_0^2} > 0$.

The marginal product of internal research, $-\frac{\partial \phi}{\partial r_0}$ could be enhanced by public science, so that $-\frac{\partial^2 \phi}{\partial r_0 \partial u} \geq 0$. Complementarity may seem natural from the perspective of absorptive capacity (Cohen & Levinthal, 1989). However, insofar as firms can innovate – introduce new products in the market – by licensing external inventions or acquiring university startups, public science may *substitute* for internal research so that $-\frac{\partial^2 \phi}{\partial r_0 \partial u} < 0$.

In Stage 1, firms may invest in research. Firm 0's research investment is denoted by r_0, and the cost of research is modelled simply as $\frac{\gamma}{2}r_0^2$, so that the value of the firm, $v_0 = kd_0 - \frac{c_{00}}{2}d_0^2 - bd_1 - \frac{c_{11}}{2}d_1^2 + c_{01}d_1d_0 - \phi(r_0,u)d_0 - \frac{\gamma}{2}r_0^2$. The value of the rival is $v_1 = d_1 - \frac{c_{00}}{2}d_1^2 - bd_0 - \frac{c_{11}}{2}d_0^2 + c_{01}d_1d_0 - \phi(r_1,u)d_1 - \frac{\gamma}{2}r_1^2$. The value functions are symmetric with one exception, $k > 1$, which captures the higher marginal payoff from innovation for the leader. The parameter c_{01} captures the nature of strategic interactions. It is positive if innovations are *strategic complements*, and negative if innovations are *strategic substitutes*.[16]

3.1 Empirical Implications

We provide details and proofs in Appendix C. Here we provide the main results and the intuition. Although the profit and cost functions are otherwise symmetric, the higher marginal payoff from innovation for the leader can result in markedly different outcomes. The returns to investing in research depend on the scale of innovation because research reduces the unit cost of innovation. In equilibrium, Firm 0 innovates more, and thus will have a higher marginal return from research. If innovations of the leader and follower are strategic substitutes, an increase in innovation by the leader reduces innovation by the follower, and thereby also reduces its incentives to invest in research. In other words, the difference between the leader and follower in the value of innovation leads to a corresponding divergence in the marginal returns to investing in research if innovations are strategic substitutes. Indeed, if the difference is large, the follower may not invest in research at all.

In what follows, we focus on the case where only the leader invests in research. Appendix Table C2 describes firm incentives to conduct research. As k increases,

---

[16] Formally, innovations are strategic complements if innovation by a firm increases the marginal payoff from innovation of its rivals, and strategic substitutes if the marginal payoff from innovation of the rival falls.

the marginal return to the leader (Firm 0) of investing in research also increases. The first result in Table C2 is that investments in science are increasing in the firm's proximity to the technological frontier (k).

Public science lowers the cost of innovation. The resulting increase in innovation raises the marginal return to investment in research. There is a second effect that depends on the relationship between public science and private research in reducing the cost of innovation. If public science complements internal research, it will increase the marginal returns to internal research, and decrease it otherwise. Finally, there is a third effect which depends on strategic interactions. If innovations are strategic substitutes, public science will increase the marginal return to internal research for the focal firm by suppressing innovation by the rival. In other words, if public science increases the marginal product of internal research in reducing the cost of innovation, and if innovations are strategic substitutes, then public science will also increase internal research. In other cases, the net effect remains an empirical matter.

Crucially, as k increases, the leader's marginal returns to investing in research will depend on the supply of public science: the supply of public science will enhance the effect of technological leadership on research if it increases the marginal product of research. Conversely, if public science and corporate research are substitutes, then public science will dampen the effect of leadership on internal research. The third result, in Table C2, therefore, is that as the leader moves closer to the technological frontier, its research investments will rise with a decrease in public science only if public science and internal research are substitutes (i.e., if public science reduces the marginal productivity of internal research).

4 Data

Our unbalanced panel of firms combines financial statements data for public (listed) and (large) private firms from Moody's Manuals; market value data for listed companies from CRSP; United States Patent Office (USPTO) data from Google Patents; and publication, lab, and scientist data from Microsoft Academic Graph (MAG), the IRL and AMS directories respectively. The combined dataset covers the period 1926-1940.

We begin with 231 B&M firms (major "industrial", i.e., non-financial firms) and their subsidiaries that patent at least once between 1926 and 1940 in an IPC that cites at least five scientific articles between 1947 and 1957.[17] The B&M sample of industrial firms collectively accounts for more than half of all non-financial corporate assets in America in the 1930s (Kandel et al., 2019). We focus on patenting firms, restricting the initial sample to firms that are "at risk" of engaging in scientific research.[18] We augment the B&M sample with 235 additional listed firms from CRSP that patent in science-citing IPCs (Kogan, Papanikolaou, Seru, & Stoffman, 2017).[19] Our basic sample thus consists of 466 private and public American firms that patent at least once between 1926 and 1940 in an IPC that cites at least five scientific articles between 1947 and 1957. Of these, there are 4,282 firm-years for which we have financial statement data between 1926 and 1940.

---

[17] Subsidiaries data are from Kandel et al. (2019) whose source is Moody's Manuals. As noted above, the Manuals include balance sheet data on important, even if unlisted, firms. For instance, Ford Motor Company, whose Initial Public Offering was only in 1956 (after the end of our sample period) has its assets and sales data reported in Moody's.

[18] Examples of excluded patent classes include B27M (woodworking), B60P (loading transportation vehicles) and E03D (Water Closets or Flushing Valves thereof). Around 26% of patenting firms are lost due to this restriction.

[19] This takes into account entry of new firms after the B&M list was assembled (in 1932). It also covers inventing firms that were not large enough to be included in the B&M sample.

4.1 Corporate Investment in Science

Scientific Publications — We source 283,992 peer-reviewed scientific papers (excluding humanities and the social sciences) in Microsoft Academic Graph (MAG) published between 1926 and 1940.[20] We match 140,766 author affiliations from these publications to our sample firm names using a fuzzy string-matching algorithm that takes into account abbreviations common in the era (e.g., firms in the railroad sector may be abbreviated as RR (railroad), RW (railway), RC (rail company)), and name variants for certain companies (e.g., AT&T's Bell Labs, SOCONY for the Standard Oil Company of New York). We ensure that articles authored by eponymous charitable foundations and hospitals (e.g., by DuPont, Carnegie and Rockefeller) are not erroneously classified as corporate publications. We match 3,263 corporate publications to 201 sample firms. Of these, 110 firms are found in the B&M sample, 162 are found in CRSP and 71 are found in both samples.[21]

Corporate Labs — We obtain data on R&D labs operated by firms from a national survey by the National Research Council (NRC) conducted since 1920 (National Research Council, 1931). Data from these surveys have been used in Mowery and Rosenberg (1998), Nicholas (2011), Field (2003) and Furman and MacGarvie (2007). We manually search the names of our firms in the entries of the 1927 (999 firms), 1931 (1,620 firms), 1933 (1,562 firms), 1938 (1,769 firms) and 1940 (2,264 firms) surveys. Some of the top publishers such as AT&T and GE, also operate the

---

[20] We extend this match to publications between 1900 and 1925 for the "pre-period" analyses in Figure 2 and Appendix Table D1. See https://www.oecd.org/science/inno/38235147.pdf for a list of the fields of science, based on the 2002 revision (6th edition) of the Frascati Manual. The Manual lays out the OECD standard for classifying scientific activity by research area and has been maintained by the National Experts on Science and Technology Indicators (NESTI) Group at the OECD since 1963.

[21] See Appendix A.1 for details on matching scientific publications to firms.

largest labs, but the rank correspondence is not one-to-one. For instance, DuPont operates a large lab with over 3,096 personnel, but publishes a total of 21 papers in our sample period. This may reflect a heterogeneity in firm publishing policies, as well as field-specific publishing behavior.

American Men of Science — We use the employment of prominent scientists by a firm as an additional indicator of corporate investment in science. For this, we use information on the employment affiliation of American scientists from the Cattel Directory of American Men of Science.[22]

The three indicators of corporate investment in science are highly correlated.[23] However, each measure is imperfect. Table 1 shows that nearly a half of all firms that operate research labs do not publish in the sciences, and only around a third employ prominent scientists that appear in the AMS directory. Historical accounts also indicate that some labs were engaged exclusively in downstream development rather than scientific research. We therefore use multiple indicators in the empirical analysis.

4.2 University Science

4.2.1 Gap in University Science

We proxy the gap between the science needed by industry and that available from American universities by measuring the "relative backwardness" of American academia compared to Europe by scientific field. Our primary measure is based on the citation-weighted number of scientific publications authored by scientists in

---

[22] The American Men of Science (AMS) directory is a comprehensive listing of prominent scientists in the United States since 1906 (Moser et al., 2022).

[23] Pairwise correlations are 0.562 between lab size and publications, 0.656 between lab size and AMS scientists, and 0.684 between publications and AMS scientists. This is despite the fact that the AMS measure is based on earlier (1921) data.

each region. We find similar results using an alternative measure based on citations to European journals made by American scientific journals. These gaps were meaningful for firms requiring frontier science during our sample period because knowledge flows from Europe were restricted after the onset of World War I.[24]

Scientific Publications: U.S. and Europe — We use the country of correspondence for the authors of scientific publications. We first collect the address information of authors for 44,355 scientific papers published between 1900 and 1920 from Clarivate Web of Science (WoS) and classify addresses into US, Europe and "Rest of World" regions based on their country names.[25] For publications with missing addresses, we match the authors' last and first names to the *American Men of Science* directory to identify 27,924 publications by prominent American scientists. The rest of the publications during this period are classified as European. We exclude papers in the social sciences and humanities and are left with 12 OECD subfields for which at least one "European" or "American" scientist published between 1900 and 1920.[26] The above process yields 155,571 publications by Europeans and 60,605 publications by Americans in the sciences between 1900 and 1920. We weigh the publication counts by the number of forward paper citations received until 2019.

Constructing Scientific Gaps for Firms — Our regional scientific activity data from Web of Science are encoded at the scientific field level. Therefore, we link

---

[24] Iaria et al. (2018) show that scientific knowledge transfer from Europe was disrupted since the onset of World War I and failed to recover as late as 1930. Consistent with a high cost of accessing German knowledge after World War I, Moser et al. (2014) show that Jewish chemists fleeing persecution from the Third Reich changed the technological direction of American chemical patents.

[25] We use Clarivate Web of Science because Microsoft Academic Graph does not contain a separate address field.

[26] We use the correspondence in Marx and Fuegi (2020) to map Web of Science subject fields to 39 OECD subfields. Appendix Table B1 provides the breakdown by field.

them to firms based on how much the firm patents in a patent class, and on how much the patent class relies on a scientific field. We first calculate the number of European papers published between 1900 and 1920 relevant to a patent class by weighting the number of European papers in each field by how often they are cited by patents in the class.[27] We sum the weighted papers over all scientific fields at the patent class level to produce European papers relevant for each class. We then weight these papers by how often the focal firm patents in the class.[28] We sum the weighted papers over all patent classes to generate the number of European papers relevant to the firm. The number of American publications relevant to the firm is obtained analogously. We then divide the number of European publications by the sum of American and European publications to obtain our primary measure of scientific gap the firm faces.[29]

4.2.2 Geographical Distance to Universities

We supplement our measure of the supply of knowledge from American universities (measured through the aforementioned comparisons to Europe) by calculating firms' geographical proximity to universities, which could affect the access a firm

---

[27] The weights divide the number of patent citations made to the field by total patent citations to science from the IPC between 1947 and 1957 (the first 10-year period from the time NPL citations were formalized in U.S. patent documents).

[28] The weights divide the number of patents issued to the firm in the IPC by total firm patents between 1926 and 1940.

[29] For example, 15% of AT&T's patents granted between 1926 and 1940 are in IPC H01J (Electric discharge tubes or discharge lamps). Patents in this IPC, in turn, cite the Chemical Sciences most often (26%), followed by Electrical Engineering (23%) and Physical Sciences (21%) between 1947 and 1957. As we see in Appendix Table B1, Chemical Sciences and Physical Sciences have European-to-American ratios that are higher than the average, which contributes to the high (in the 90th percentile) firm-level gap score for AT&T. In contrast, General Ice Cream Corp, which is below the 10th percentile in this gap score, patents most often in A23G (Cocoa; Cocoa Products), where the highest number of NPL citations are made to Biological Sciences. Biological sciences, in turn, has a European-to-American ratio below the average, which contributes to the firm receiving a low gap score. Appendix B presents a detailed example, showing each step of the calculation.

has to human capital. For each sample firm we calculate the average distance between its headquarters' location and all American universities granting graduate degrees in the natural sciences in 1930. We find the headquarters' locations from the 1929-1930 edition of Moody's Manuals. For firms with no address information, we match their patents to the HistPat dataset, which extracts Federal Information Processing System (FIPS) County codes for patent assignees (Petralia, Balland, & Rigby, 2016). We impute the firm's address as the one that appears most frequently in its patents.

For university addresses, we use the 1930 edition of the "List of American Doctoral Dissertations", a catalog published by the Library of Congress, to identify 41 American universities that granted graduate degrees in the natural sciences during our sample period. We manually search for the addresses of these institutions and calculate their geodesic distances to our sample firms.[30]

4.3 Patents

Patent data are derived from Google's public patent dataset. There are 637,190 patents granted between 1926 and 1940 by the USPTO. We normalize forward patent citations by dividing the total number of prior-art citations received by a focal patent by the per-patent citations received by all patents granted in the focal patents' issue year. To measure the extent to which a patent "relies" on science, we count citations to scientific publications in Microsoft Academic Graph in the text of the patent, from Marx and Fuegi (2020).[31] We measure the "novelty" of a

---

[30] Appendix A.2 provides further details.

[31] We use in-text citations because NPL citations are available only after 1947. We use references with a confidence score above 8. We find 237 patent citations to science by our sample firms between 1926 and 1940.

patent by measuring whether it is the first in its Cooperative Classification Class (CPC). Finally, we use the patent "importance" measure developed by Kelly et al. (2021). This measure identifies patents distinct from previous work but related to subsequent innovations by dividing a patent's 10-year forward textual similarity by its 5-year backward similarity. Appendix A.1 describes how patent assignees are matched to firms using the same fuzzy string-matching algorithm used to match publications. We match 92,330 patents to the 466 firms in our panel between 1926 and 1940.

4.4 Financial Data and Industry Concentration

Financial Statement Variables — Balance sheet and earnings data are not available before 1950 from conventional sources such as S&P Compustat. Therefore, we build on Kandel et al. (2019), who collect data on firm assets and earnings for the sample firms for the years 1926, 1929, 1932, 1937 and 1940 using Moody's Manuals, and fill in data for the intervening years from the same source.[32] We classify firms to industries by matching descriptions of firm "occupations" in Moody's Manuals by hand to one of the 85 3-digit industry codes in the revised 1947 SIC tables (reported by the BEA in 1958 (Department of Commerce, 1965)). We augment the dataset with end-of-the-year stock market value data for all listed firms using the CRSP Monthly Stock File for North American firms. For listed firms that appear on CRSP but not in the B&M sample, we obtain financial data from Graham et al. (2015), who manually collected the data from Moody's Industrial Manuals.

Market Share and Industry Concentration — We measure annual market share

---

[32] We use additional, ownership-related variables, drawn from Moody's Manuals in a robustness test in Appendix Table D2

by dividing firm sales in a year by 3-digit industry sales in the same year.[33] We then average annual market share for the first five years of the panel (1926-1930) for each firm.[34] We also use Wilcox (1940) to classify 3-digit industries into monopolistic, oligopolistic and competitive ones. Wilcox uses a broad set of criteria, as well as the regional nature of the markets, to separate effectively competitive industries from effectively monopolistic ones between 1934 and 1939 (the measures do not refer explicitly to a specific year). In our analysis, we focus on the distinction between competitive and non-competitive industries.[35]

4.5 Descriptive Statistics

The maximal number of observations is 6,990 (466 firms observed over 15 years between 1926 and 1940).[36] "Lab Employees" counts the number of lab personnel reported in the IRL directory. There are only around a third of the total observations for "Lab Employees" because the IRL was collected for only five years of our sample period by the NRC (1927, 31, 33, 37, 40). The difference in median and mean values indicates scientific publications and lab personnel are skewed to the right.

Section 2 above documents a rise in aggregate corporate investment in science during the interwar period. However, this masked substantial heterogeneity across firms. Table 1 shows that 154 firms out of our sample of 466 firms never operate an R&D lab, while more than half (265) of the firms never publish a scientific article. Perazich and Field (1940) independently estimate that fewer than 1% of all firms

---

[33] We use gross income to measure sales as our sample consists primarily of firms in manufacturing.

[34] We use earlier market share data mitigate concerns that investment in science may affect market share later in the panel. Averaging mitigates the potential bias caused by years with incomplete sales data.

[35] Stigler (1949) and Nutter and Einhorn (1969) both validate the Wilcox (1940) classification.

[36] Appendix Table A2 presents descriptive statistics at the firm-year level. A corresponding variable definition table can be found in Appendix Table A1.

accounted for a third of all industrial research employment in 1921, 1927 and 1938; A mere 45 firms in 1938 employed half of the total research personnel.

Figure 3 shows that the unconditional probability of a firm having a scientific patent, one of our indicators of the firm having reached the scientific frontier, is positively related to the firm's investment in science. Firms with scientific patents are more likely i) to have a lab, ii) to publish, or iii) to employ prominent scientists. This relationship is stronger when we combine indicators for corporate investment in science. For instance, 1.8% of firms that engage in one of the aforementioned activities (the "1 Indicator" group) cite scientific articles, which is larger than those that engage in none of them (the "No Science" group, with 0.9%). However, this difference is small and statistically insignificant. In contrast, the 27.6% of firms that engage in all three activities cite scientific articles in their patents, and the difference with the rest of the groups is statistically significant. The results are qualitatively similar for firms that engage in all three activities with respect to other patent-based measures of proximity to the technological frontier, as well as for measures of leadership (market share), size (assets) and operation in a concentrated market (although the patterns for this industry-level measure are less pronounced).

Table 2 shows that the three measures of proximity to the technological frontier are positively correlated, though not perfectly. This implies each measure captures a different technological characteristic of the firm. We therefore present in Table 3 specifications that include each of these measures separately, and all three of them together, to explore the relative importance of each in accounting for investment in science. Similarly, Table 4 presents regression specifications that include market share, assets and concentration separately, as well as all of them together.

5 Empirical Results

5.1 Who Invests in Science?

Table 3 presents the results of conditional Poisson specifications regressing corporate investment in science on measures of proximity to the technological frontier, controlling for firm size (assets) and geographical distance to universities (which turns out not to be significant in any specification).[37] We include year and industry dummies except for Columns 9 through 12, where the dependent variable (corporate scientists from the 1921 edition of AMS) is based on a cross-section at the firm level rather than a firm-year panel.[38]

All three indicators of investment in science (publications, lab size, and the number of AMS employed) are positively associated with patent citations to science, first-in-CPC dummy, and patent importance, both when each is included separately in the regression, and when all measures are included together (Columns 4, 8 and 12).[39] The estimates from Columns 1, 5, and 9 imply that citing scientific articles in patents is associated with nine times more corporate publications, 2.9 times more lab employees and 6.7 times more AMS scientists. Because a substan-

---

[37] Furman and MacGarvie (2007) argue that early pharmaceutical laboratories benefited from proximity to universities. George Davis, for instance, recruited two scientists from the University of Michigan for the laboratory of his firm Parke-Davis (located nearby in Detroit) in an effort to mass-produce an antitoxin for diphtheria that was discovered by German and French scientists in 1894. We find in Appendix Figure A6 that laboratory employees are indeed negatively correlated with geographical proximity to universities (top right panel), but this relationship is less pronounced for corporate publications (top left) and AMS Scientists (bottom left panel).

[38] Firm size (assets) in Columns 9 through 12 are averaged for the first five years of the sample (1926-1930) to be as close as possible to 1921, when the AMS data are collected.

[39] Given that patent cites to science are measured at the firm level, standard errors are clustered at the firm. We find similar results with a between-firm specifications in Appendix Table D2, Columns 11 and 12, when replacing the dependent variables with their firm-level averages over the sample period.

tial fraction of firms does not invest in science, we also test whether proximity to the technological frontier can predict the decision to invest in science at all. Appendix Table D2 Columns 1-3 estimate Probit specifications which show that firms with scientific patents are around 60% more likely to publish (Column 1) and 40% more likely to operate labs (Column 2). Employment of AMS scientists is not correlated with patent citations of science, but it is positively correlated with the other two measures of proximity to the scientific frontier (Column 3). The use of 1921 AMS scientists as a dependent variable in this context is based on the premise of cross-sectional (time-invariant) differences between firms; it cannot, of course, capture the evolution of corporate investment in science over time.

Given the positive correlation between corporate science and size (assets), we also examine if our results hold for the largest firms in the sample, and when the largest firms are excluded. Appendix Table D2, Column 7 indicates that, for the largest quintile of firms in the sample the magnitude of the coefficient on patent citation to science is around 30% smaller than in the full sample, but it is still statistically significant (Columns 8 and 9 show similar results for lab personnel and AMS scientists). In Column 10 we exclude eleven firms above the 95th percentile of total papers and the 95th percentile of assets and still find a positive and significant relationship between publications and measures of proximity to the technological frontier.

Table 4 focuses on the relationship of corporate science with absolute size (assets), relative size (market share) and industry concentration (competition).[40] All

---

[40] Since these measures are correlated with technological frontier measures (Table 2), we include them altogether in Appendix Table D3 and find the main results hold. In unreported robustness checks, we control for liquidity (sales normalized by assets) and find that the positive coefficients on size (assets) in Table 4 remain. Note that we measure assets (in Columns 1,5 and 9) and market

three are correlated significantly with corporate investment in science. For example, a one standard deviation higher market share is associated with 2.7 times more publications, around 4% more lab employees (though this relationship is not statistically significant) and 45% more AMS scientists. The relationship with market leadership is consistent with the theoretical framework presented above; the estimates on firm size and competition indicate the importance of the ability to appropriate the benefits of investment in science.

5.2 The Relative Backwardness of American University Science

We explore the extent to which the weakness of American academia in certain fields (i.e., the "gap" in public science) accentuates the incentives of firms to invest in research. Table 5 displays results from a conditional Poisson regression where we include a measure of the gap along with controls for size, distance to academic institutions along with year and industry dummies. The point estimate in Column 1 of Table 5 implies that a one standard deviation higher gap in university science is associated with around twice as many corporate publications. The results are qualitatively similar for lab employees and AMS scientists in Columns 2 and 3.

Table 6 interacts the university gaps with proximity to the technological frontier.[41] The interaction coefficient estimate in Column 1 suggests that the positive association between the scientific gap and investment in corporate research is driven by firms whose patents cite science (the two other measures of proximity to the frontier yield similar results). Comparing university gap measures at the 25th and 75th

---

share (in Columns 2,6 and 10) for the first five years of the panel to mitigate the possibility that earlier investments in science increase the (absolute or relative) size of firms later in the sample.

[41] Since the gap measure is a ratio between European and American publications, we also control for the level of European and American publications (unreported in the tables), whose coefficients are insignificant. The results are not sensitive to their inclusion.

percentile, firms with scientific patents have eight-fold more scientific publications, whereas firm that do not cite science exhibit a negligible change (a statistically insignificant 4% decrease).

Table 7 interacts university gaps with firm size, market share and concentration. The interaction term estimate in Column 2 implies that the association between scientific publications and the university gap is five times as high for firms with high compared to low market shares (75th percentile vs 25th), while Column 3 shows that the positive association between scientific publications and the university gap is principally for firms in non-competitive industries.[42] We see a similar pattern of results using lab employees (Columns 5 and 6) as well as AMS scientists (Column 8 and 9). In sum, the firms most likely to respond to gaps in science were technologically advanced and faced limited competition.[43]

5.3 Corporate Science and Firm Performance

If firms invest in research to solve pressing technological problems, such firms ought also to have produced better and more valuable new inventions. We feature two measures: patents deemed valuable by investors (from Kogan et al., 2017) and highly cited patents.[44] We estimate a conditional Poisson specification that regresses the number of these "home-run" patents against investment in science and controls for size (assets) as well as patent stock.[45]

---

[42] We find the interaction term with respect to size (assets) in Columns 1, 4 and 7 are imprecise, which may indicate that market share and competition capture appropriability from scientific knowledge that is distinct from size.

[43] Appendix Tables D4, D5, and D6, replicate these results with an alternative measure of university gap based on the number of citations from American journals to European ones.

[44] Patent values are available only for public firms. Public firms account for the majority of firms in our sample and constitute the vast majority (around 80%) of firms that publish or operate labs.

[45] "Home-runs" are defined as patents in the top 5% of their respective measures for each grant year. Note that patent value is conceptually distinct from patent-based measures of proximity to

The estimates from Table 8, Columns 1-4 show that two of our proxies for investment in science (publications and the number of AMS scientists, but not lab size) are positively correlated with the number of highly valuable patents. Columns 5-8 show similarly that investment in science (with the exception of lab size) is associated with the production of highly cited patents. For example, an increase in publication stock by one standard deviation is associated with about 12% more patents in the top 5% of stock market value (relative to the sample mean). Though not causal (there may be an underlying firm characteristic driving both corporate science and patenting), these results are consistent with the view that firms that invested in internal scientific research were also more likely to produce higher quality inventions.

As an alternative to estimating patent-based outcomes, we also estimate the relation between corporate science (publications) and firm stock market value for publicly traded firms through an OLS specification in Table 9, controlling for lagged asset and patent stock. A one standard deviation larger publication stock (normalized by assets) is associated with a 9% larger market to book ratio in Column 1. Column 2 shows that a standard deviation larger publication stock is associated with about 0.6% increase in logged firm market capitalization relative to the sample mean. However, splitting the sample by university gap measures in Columns 3 and 4, we find that firms whose gaps in university science are above the sample mean are the ones driving the association between market value and publication stock.[46]

---

the technological frontier used in Tables 3 and 6. Patents which cite science, are first in their technology class, and may actually be discounted by investors and future inventors *because* of their novelty.

[46] Appendix Table D7 shows a similar result using citations to European journals to measure university gap.

Columns 5 and 6 show that publication stock is positively associated with stock market value primarily for firms that employ AMS scientists and have a lab (Column 6). Firms in Column 5, which do not engage in both activities, show a negative and statistically insignificant correlation between publishing and stock market value. This is consistent with uneven quality of publications across scientific fields at that time, as well as differences in publication norms across industries. Columns 7 and 8 restrict the sample to firms that employ AMS scientists and have a lab, and indicate that the association between publication stock and stock market value is stronger for firms whose fields are related to scientific disciplines in which American universities lagged behind Europe. While not causal, the results suggest that investments in science are positively related to market value, and that this relationship is driven by firms facing significant gaps in university science.

6 Discussion and Conclusion

Technical progress is a major source of productivity increases and economic growth. Towards the beginning of the twentieth century, technical progress became increasingly reliant on scientific knowledge. Scientific discoveries, such as the vacuum tube and polymers, created opportunities for productivity advances in existing industries such as textiles, lighting, and telephony, and opened up entirely new ones, such as plastics, synthetic fibers, radio, and television. To exploit these opportunities, American companies needed a deeper understanding of the underlying science. The weakness of American academia in certain fields meant that some firms decided to invest internally in advancing them. For large firms that were close to the technological frontier and operated in concentrated industries, investment in internal research was the way to overcome this institutional weakness and thereby gain competitive advantage: competitors would not be able to readily acquire the needed scientific knowledge from universities. More generally, and in relation to public debates today, our results suggest that large, technologically advanced firms can play an important role in advancing knowledge, but that this is likely to reinforce their advantage over competitors. This may be related to why a handful of companies at the technological frontier today (chiefly in fields such as quantum computing and Artificial Intelligence) continue to invest in scientific research, while other firms do not.

How the private value from the use of a public good like scientific knowledge changes as the supply of the public good expands, and how the private value of the public good differs across firms, is an important but understudied topic. Universities produce new knowledge as well as human capital, which affect the returns to private investment in research and innovation. In the decades since the end of our sample period, American universities have increased the quality and quantity of their scientific output. Following World War II, the growth of university research was paralleled by growing investments in corporate research. However, by the 1980s, the two trends have diverged, leading to a growing division of labor between academia and universities (Arora, Belenzon, Patacconi, & Suh, 2020).[47] More research is needed to understand the implications of this division of labor in innovation for market leaders and followers, and for the rate and direction of technological advance more broadly.

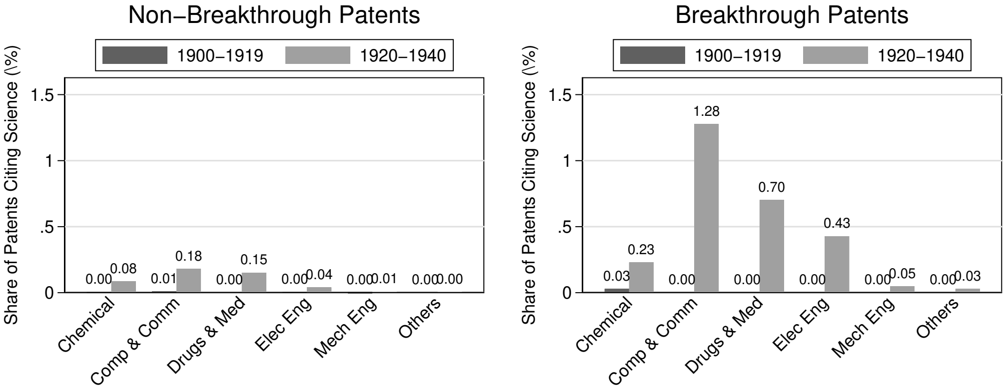
In addition to providing new evidence on corporate research in America in the

---

[47] Corporate research in the 1930s was possibly easier to protect from rivals (few could effectively use it), whereas by the 1980s knowledge spillovers may have become more costly. Indeed, Arora et al. (2021) find that companies cut back on research when spillovers to rivals increase relative to internal use.

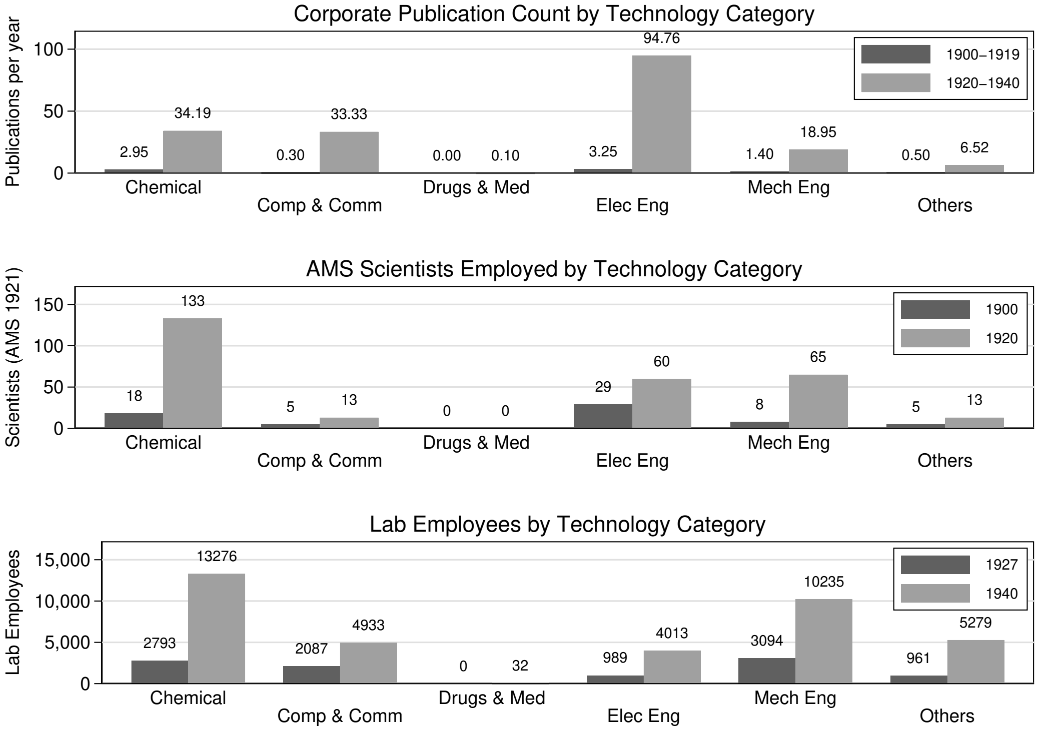
interwar period, we assemble the most extensive historical sample of American firms that were involved in innovation during that period, including information on the scientific investment of these firms and on the relative gap between American and European universities. We hope that these newly developed data will contribute to future research on the open questions we raised.

Figure 1: Patent Citations to Science, by Breakthrough Patent Status (KPST) and Technology Category



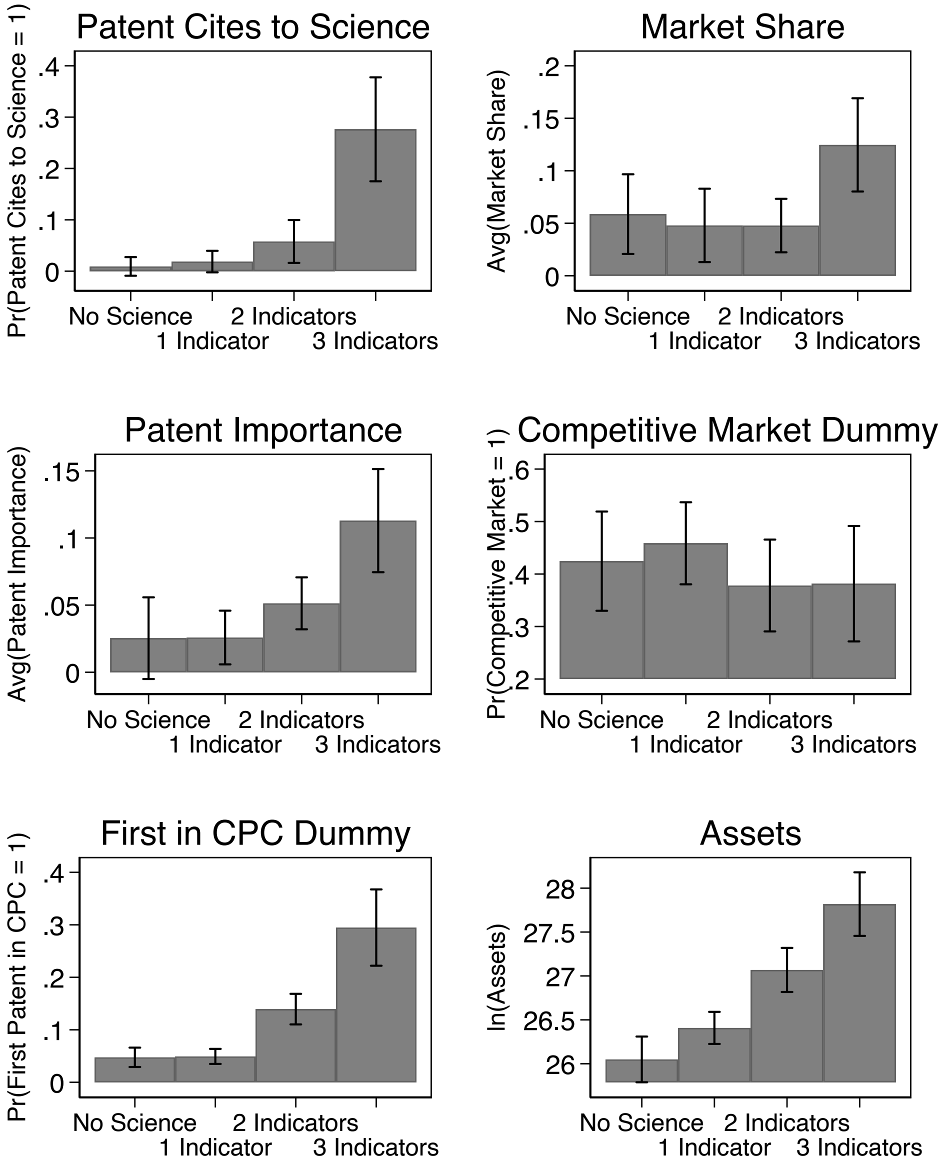
Notes: The right panel plots the percentage share of "breakthrough" patents that cite scientific articles for patents issued between 1900 and 1919 and 1920 and 1940, by NBER categories. The left panel plots the same for all other ("non-breakthrough") patents. Breakthrough patents are defined as those that are in the top 10% of the importance measures from Kelly, Papanikolaou, Seru, and Taddy (2021), where importance is calculated by dividing similarity with future patents (ten years after focal patent issuance) by similarity with past patents (five years before focal patent issuance) net of focal patent grant year fixed effects. Scientific citation data for patents is sourced from Marx and Fuegi (2020). NBER categories for patents are sourced from Hall et al. (2001).

Figure 2: CORPORATE INVESTMENTS IN SCIENCE, BY TECHNOLOGY CATEGORY



Notes: The top panel plots total corporate publications in each NBER patent category per year produced between 1900 and 1919 and between 1920 and 1940 respectively. The middle panel plots the number of scientists from the AMS directory affiliated with firms in each category in 1900 and 1920. The bottom panel plots total employees at industrial laboratories in each category per year for 1927 and 1940. A technology category for a firm is defined as the NBER patent category in which it is granted the most patents during our main sample period between 1926 and 1940.

Figure 3: INDICATORS OF CORPORATE SCIENCE AND DISTANCE TO TECHNOLOGICAL FRONTIER AND FIRM CHARACTERISTICS



Notes: The left panel plots measures of technological leadership by investment in science while right panel plots market share, size (assets) and concentration (competition dummy) analogously. The three indicators of investment in science on the x-axis measure whether the firm i) produces a scientific publication; ii) operates an industrial laboratory; iii) employs a scientist from the AMS directory. The “No Science” group consists of firms that engage in none of these three activities. The “3 Indicators” firms engage in all three activities. The “2 Indicators” firms engage in two of the three activities. “1 Indicator” firms engage in one of the three activities. “Market Share” and “Assets” are averaged between 1926 and 1930.

Table 1: Cross Tabulation of Measures of Corporate Science

	Does Not Publish	Publishes	Publishing Firm Share
Does Not Operate Lab	113	41	27%
Operates Lab	152	160	51%
	Does Not Employ AMS	Employs AMS	AMS Firm Share
Does Not Operate Lab	138	16	10%
Operates Lab	206	106	34%
	Does Not Employ AMS	Employs AMS	AMS Firm Share
Does not publish	228	37	14%
Publishes	116	85	42%

Notes: The unit of analysis is the firm. The top and middle panels split the sample by whether a firm operates a research lab during our sample period based on the IRL directory. The bottom panel splits the sample by whether the firm produces a scientific publication. The columns in the top panel further split the sample by whether the firm produces a scientific publication; those in the middle and bottom panels split the sample by whether the firm employs scientists from the 1921 edition of AMS.

Table 2: Correlations Between Explanatory Variables

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Dummy for Patent Cite to Science	1.000					
(2) Dummy for First Patent in CPC	0.591***	1.000				
(3) Patent Importance	0.280***	0.267***	1.000			
(4) ln(Assets, 1926-1930)	0.338***	0.396***	0.128*	1.000		
(5) Market Share	0.180**	0.294***	-0.004	0.413***	1.000	
(6) Dummy for Competitive Market	0.071	-0.085	0.107	0.048	0.093	1.000

Notes: This table displays pairwise Pearson correlations for the main explanatory variables relating to technological and market leadership at the firm level. $^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$

Table 3: CORPORATE INVESTMENT IN SCIENCE & TECHNOLOGICAL LEADERSHIP (POISSON)

Dependent Variable	Publication Count				Lab Employees				AMS Scientists			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dummy for Patent Cite to Science	2.215			1.444	1.063			0.698	1.939			1.335
	(0.437)			(0.454)	(0.294)			(0.300)	(0.350)			(0.362)
Dummy for First Patent in CPC		1.851		1.110		1.033		0.693		2.617		1.889
		(0.269)		(0.230)		(0.175)		(0.159)		(0.551)		(0.405)
Patent Importance			5.990	2.952			1.160	0.772			4.157	1.045
			(1.383)	(1.051)			(0.443)	(0.310)			(0.890)	(0.888)
Distance to Universities	0.000	-0.000	-0.001	0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.001	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
ln(Assets)	0.629	0.677	0.877	0.561	0.707	0.670	0.771	0.588	0.615	0.584	0.849	0.455
	(0.205)	(0.215)	(0.218)	(0.147)	(0.074)	(0.076)	(0.090)	(0.070)	(0.152)	(0.190)	(0.164)	(0.163)
Average of Dep Var	0.688	0.688	0.939	0.939	61.267	61.267	82.962	82.962	2.261	2.261	2.886	2.886
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.670	0.649	0.665	0.725	0.691	0.692	0.668	0.716	0.630	0.645	0.585	0.704
Number of Firms	422	422	397	397	394	394	338	338	276	276	202	202
Number of Obs	3,855	3,855	2,640	2,640	1,331	1,331	934	934	276	276	202	202

Notes: The analysis is at the firm-year level for Columns 1-8 and at the firm level in Columns 9-12. "ln(Assets)" takes the natural log of concurrent assets for Columns 1-9, and of average assets between 1926 and 1930 for Columns 9-12. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 4: Corporate Investment in Science & Market Leadership (Poisson)

Dependent Variable	Publication Count				Lab Employees				AMS Scientists			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ln(Assets, 1926-1930)	1.180			0.796	0.927			0.857	0.985	0.783	0.807	0.669
	(0.262)			(0.244)	(0.108)			(0.149)	(0.167)	(0.208)	(0.089)	(0.100)
Market Share		6.634		2.209		0.260		0.254		2.470		1.676
		(0.790)		(1.463)		(0.941)		(0.844)		(1.109)		(0.948)
Dummy for Competitive Market			-1.076	-1.019			-0.472	-0.555			-0.712	-0.827
			(0.521)	(0.509)			(0.235)	(0.285)			(0.354)	(0.379)
Distance to Universities	-0.001	-0.001	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
ln(Assets)		0.555	0.892			0.860	0.846					
		(0.104)	(0.113)			(0.137)	(0.074)					
Average of Dep Var	0.628	0.922	0.624	0.716	56.157	78.168	58.817	65.549	2.261	2.789	1.768	2.184
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No
3-Digit SIC Dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No
Pseudo-R^2	0.592	0.722	0.320	0.379	0.623	0.678	0.495	0.431	0.517	0.555	0.324	0.358
Number of Firms	355	273	417	272	355	260	389	272	276	209	354	272
Number of Obs	4,860	2,747	4,252	4,080	1,720	934	1,386	1,360	276	209	354	272

Notes: The analysis is at the firm-year level for Columns 1-8 and at the firm level in Columns 9-12. “ln(Assets)” takes the natural log of concurrent assets; ln(Assets, 1926-1930) takes the average assets between 1926 and 1930. “Market Share” averages firm market share at the 3-digit industry level between 1926 and 1930. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 5: CORPORATE INVESTMENT IN SCIENCE AND GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	Publication Count	Lab Employees	AMS Scientists
	(1)	(2)	(3)
University Gap	22.559	10.332	10.932
	(8.351)	(4.853)	(7.104)
ln(Assets)	0.992	0.831	0.963
	(0.231)	(0.075)	(0.149)
Distance to Universities	-0.001	-0.001	-0.001
	(0.001)	(0.000)	(0.001)
Average of Dependent Variable	0.688	61.267	2.261
Year Dummies	Yes	Yes	No
3-Digit SIC Dummies	Yes	Yes	Yes
Pseudo-R^2	0.632	0.668	0.547
Number of Firms	422	394	276
Number of Observations	3,855	1,331	276

Notes: The analysis is at the firm-year level for Columns 1 and 2 and at the firm level for Column 3. "ln(Assets)" takes the natural log of concurrent assets for Columns 1 and 2, and of average assets between 1926 and 1930 for Column 3. "University Gap" divides the number of European scientific publications relevant to a firm from Web of Science between 1900 and 1920 by all (European and American) publications relevant to it. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 6: Corporate Investment in Science and Gap in University Science, by Technological Leadership (Poisson)

Dependent Variable	Publication Count			Lab Employees			AMS Scientists		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
University Gap	-1.637	-1.438	19.055	7.428	4.736	17.252	-6.589	-18.326	-0.215
	(7.162)	(8.656)	(8.482)	(4.832)	(4.985)	(5.194)	(5.329)	(8.061)	(7.676)
× Dummy for Patent Cite to Science	86.755			18.209			41.809		
	(14.398)			(12.098)			(11.088)		
× Dummy for First Patent in CPC		55.629			18.150			43.570	
		(11.037)			(8.127)			(15.468)	
× Patent Importance			164.748			50.718			114.024
			(25.289)			(23.017)			(65.463)
Dummy for Patent Cite to Science	-59.282			-11.725			-27.314		
	(10.057)			(8.389)			(7.795)		
Dummy for First Patent in CPC		-37.708			-11.743			-28.083	
		(7.762)			(5.642)			(11.095)	
Patent Importance			-111.665			-34.301			-76.495
			(17.557)			(15.813)			(45.457)
ln(Assets)	0.582	0.640	0.926	0.686	0.643	0.748	0.549	0.530	0.808
	(0.100)	(0.128)	(0.118)	(0.077)	(0.075)	(0.070)	(0.152)	(0.178)	(0.162)
Distance to Universities	0.000	0.000	-0.001	-0.000	-0.000	-0.000	-0.000	0.000	-0.001
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Average of Dependent Variable	0.688	0.688	0.939	61.267	61.267	82.962	2.261	2.261	2.886
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.734	0.702	0.722	0.700	0.705	0.690	0.657	0.678	0.617
Number of Firms	422	422	397	394	394	338	276	276	202
Number of Observations	3,855	3,855	2,640	1,331	1,331	934	276	276	202

Notes: The analysis is at the firm-year level for Columns 1-6 and at the firm level for Columns 7-9. “ln(Assets)” takes the natural log of concurrent assets for Columns 1-6, and of average assets between 1926-1930 for Columns 7-9. “University Gap” divides the number of European scientific publications relevant to a firm from Web of Science between 1900 and 1920 by all (European and American) publications relevant to it. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 7: CORPORATE INVESTMENT IN SCIENCE AND GAP IN UNIVERSITY SCIENCE, BY MARKET LEADERSHIP (POISSON)

Dependent Variable	Publication Count			Lab Employees			AMS Scientists		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
University Gap	-223.086	-0.752	21.979	-32.146	4.810	20.323	88.239	-4.469	14.240
	(215.095)	(7.485)	(6.290)	(110.641)	(6.777)	(4.212)	(140.315)	(7.601)	(5.716)
× ln(Assets, 1926-1930)	8.782			1.358			-2.786		
	(7.773)			(4.075)			(5.155)		
× Market Share		123.761			82.924			104.053	
		(48.402)			(34.074)			(28.248)	
× Dummy for Competitive Market			-23.515			-15.090			-1.941
			(7.977)			(6.132)			(11.151)
ln(Assets, 1926-1930)	-5.188			-0.061			2.954	0.743	0.832
	(5.449)			(2.871)			(3.632)	(0.196)	(0.093)
Market Share		-81.999			-58.589			-70.922	
		(34.763)			(23.881)			(19.882)	
Dummy for Competitive Market			15.668			9.970			0.493
			(5.707)			(4.312)			(7.760)
Distance to Universities	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
ln(Assets)		0.596	0.839		0.856	0.841			
		(0.071)	(0.063)		(0.124)	(0.062)			
Average of Dependent Variable	0.628	0.922	0.624	56.157	78.168	58.817	2.261	2.789	1.768
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
3-Digit SIC Dummies	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Pseudo-R^2	0.634	0.739	0.393	0.627	0.695	0.540	0.548	0.590	0.371
Number of Firms	355	273	417	355	260	389	276	209	354
Number of Observations	4,860	2,747	4,252	1,720	934	1,386	276	209	354

Notes: The analysis is at the firm-year level for Columns 1-6 and at the firm level for Columns 7-9. “ln(Assets)” takes the natural log of concurrent assets; “ln(Assets, 1926-1930)” takes average assets between 1926 and 1930. “Market Share” averages firm market share at the 3-digit industry level between 1926 and 1930. “University Gap” divides the number of European scientific publications relevant to a firm from Web of Science between 1900 and 1920 by all (European and American) publications relevant to it. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 8: CORPORATE SCIENCE AND PATENT VALUE (POISSON)

Dependent Variable	Top 5% Market Value (KPSS)				Top 5% Forward Cites			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Publication Stock)	0.242			-0.620	0.162			0.311
	(0.116)			(0.353)	(0.046)			(0.130)
ln(Lab Employees)		-0.099		-0.585		0.076		0.042
		(0.080)		(0.301)		(0.022)		(0.057)
ln(AMS Scientists)			3.211	3.993			0.277	0.046
			(0.737)	(0.842)			(0.088)	(0.118)
ln(Assets)	0.876	1.067	0.480	0.635	-0.011	-0.033	0.303	0.176
	(0.175)	(0.125)	(0.498)	(0.491)	(0.053)	(0.048)	(0.086)	(0.103)
ln(Patent Stock)	0.398	0.570	0.372	0.322	0.920	1.001	0.431	0.393
	(0.091)	(0.139)	(0.312)	(0.313)	(0.041)	(0.049)	(0.045)	(0.044)
Average of Dep Var	1.521	1.530	0.171	0.171	1.011	1.126	0.892	0.892
Year Dummies	Yes	Yes	No	No	Yes	Yes	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.795	0.820	0.833	0.838	0.703	0.700	0.647	0.665
Number of Firms	325	320	86	86	422	394	340	340
Number of Obs	2,027	670	86	86	4,213	1,323	340	340

Notes: The analysis is at the firm-year level for all columns except for Columns 3,4,7 and 8, which are firm-level specifications. "ln(Assets)", "ln(Patent Stock)", "ln(Publication Stock)" and "ln(Lab Employees)" take the natural log of concurrent assets, patent stock, publication stock and lab employees respectively for firm-year specifications, and of averages of these values between 1926 and 1930 for firm-level specifications. Patent and publication stock are calculated using a perpetual inventory method with a 15% rate of depreciation. The dependent variable for Columns 1 and 2 is the number of firm patents in the top 5% of stock market value (Kogan et al., 2017); for Columns 3 and 4 it takes the average number of such patents over the full sample period (1926-1940). The dependent variable for Columns 5 and 6 is the number of firm patents in the top 5% in terms of forward citations; for Columns 7 and 8 it takes the average number of such patents over the full sample period (1926-1940). Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table 9: CORPORATE SCIENCE AND STOCK MARKET VALUE (OLS)

Dependent Variable	ln(Tobin's Q)	ln(Market Capitalization)						
			University Gap Split		Have Both Lab & AMS Scientist?		University Gap Split for Subsample with Lab & AMS Scientist	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	University Gap Below Mean	University Gap Above Mean	No	Yes	University Gap Below Mean	University Gap Above Mean
Publication Stock/Assets$_{t-1}$	6.860							
	(4.247)							
Patent Stock/Assets$_{t-1}$	0.150							
	(0.088)							
ln(Publication Stock$_{t-1}$)		0.165	0.119	0.200	-0.052	0.202	0.135	0.299
		(0.054)	(0.129)	(0.059)	(0.085)	(0.070)	(0.131)	(0.082)
ln(Patent Stock$_{t-1}$)		0.128	0.190	0.114	0.098	0.141	0.150	0.082
		(0.032)	(0.050)	(0.037)	(0.040)	(0.049)	(0.073)	(0.044)
ln(Assets$_{t-1}$)		0.748	0.730	0.735	0.726	0.834	0.992	0.738
		(0.049)	(0.077)	(0.062)	(0.070)	(0.078)	(0.152)	(0.075)
Average of Dependent Variable	-0.479	19.418	19.348	19.490	19.041	20.463	20.357	20.573
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.350	0.731	0.717	0.805	0.665	0.851	0.841	0.925
Number of Firms	316	321	167	154	244	77	40	37
Number of Observations	3,213	3,333	1,681	1,651	2,450	883	449	434

Notes: The analysis is at the firm-year level. Dependent variable is log of Tobin's Q in Column 1 and log of market capitalization in all other columns. Columns 3 and 4 split the sample by mean values of the "University Gap" measure based on European and American publications. Columns 5 and 6 split the sample by whether the firm both employs lab personnel as well as AMS scientists. Columns 7 and 8 split the Column 6 sample (i.e., firms operating a lab and employing AMS scientists) by their average "University Gap" measure. Standard errors are clustered at the firm level.

References

Arora, A., Belenzon, S., & Patacconi, A. (2019). A Theory of the US Innovation Ecosystem: Evolution and the Social Value of Diversity. *Industrial and Corporate Change*, *28*(2), 289–307.

Arora, A., Belenzon, S., Patacconi, A., & Suh, J. (2020). The Changing Structure of American Innovation: Some Cautionary Remarks for Economic Growth. *Innovation Policy and the Economy*, *20*(1), 39–93.

Arora, A., Belenzon, S., & Sheer, L. (2021). Knowledge Spillovers and Corporate Investment in Scientific Research. *American Economic Review*, *111*(3), 871–98.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Azoulay, P. (2002). Do Pharmaceutical Sales Respond to Scientific Evidence? *Journal of Economics & Management Strategy*, *11*(4), 551–594.

Babina, T., Bernstein, A., & Mezzanotti, F. (2021). Crisis Innovation: Evidence from the Great Depression.

Berle, A., & Means, G. (1932). Private Property and the Modern Corporation. *New York: Mac-millan*.

Birr, K. (1957). *Pioneering in Industrial Research: The Story of the General Electric Research Laboratory*. Public Affairs Press.

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, *81*(4), 1347–1393.

Boroush, M. (2022, June 1). *U.S. R&D Increased by $62 Billion in 2019 to $667 Billion; Estimate for 2020 Indicates a Further Rise to $708 Billion* (InfoBrief NSF 22-330). National Center for Science and Engineering Statistics. Alexandria, VA.

Cerveaux, A. (2013). Taming the Microworld: DuPont and the Interwar Rise of Fundamental Industrial Research. *Technology and Culture*, *54*(2), 262–288.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R & D. *The Economic Journal*, *99*(397), 569–596.

David, P. A., Hall, B. H., & Toole, A. A. (2000). Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence. *Research Policy*, *29*(4-5), 497–529.

Department of Commerce. (1965, September). *Survey of Current Business* (45(9)).

Field, A. J. (2003). The Most Technologically Progressive Decade of the Century. *American Economic Review*, *93*(4), 1399–1413.

Furman, J. L., & MacGarvie, M. J. (2007). Academic Science and the Birth of Industrial Research Laboratories in the US Pharmaceutical Industry. *Journal of Economic Behavior & Organization*, *63*(4), 756–776.

Graham, J. R., Leary, M. T., & Roberts, M. R. (2015). A Century of Capital Structure: The Leveraging of Corporate America. *Journal of Financial Economics*, *118*(3), 658–683.

Gross, D. P., & Sampat, B. N. (2022). *The World War II Crisis Innovation Model: What Was It, and Where Does It Apply?* NBER Working Papers: No. 27909.

Gutiérrez, G., & Philippon, T. (2020). *Some Facts about Dominant Firms*. NBER Working Papers: No. 27985.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools*. National Bureau of Economic Research.

Hicks, D. (1995). Published Papers, Tacit Competencies and Corporate Management of the Public/Private Character of Knowledge. *Industrial and Corporate Change*, *4*(2), 401–424.

Hoddeson, L. (1981). The Emergence of Basic Research in the Bell Telephone System, 1875-1915. *Technology and Culture*, *22*(3), 512–544.

Hounshell, D., & Smith, J. K. J. (1988). *Science and Corporate Strategy : Du Pont R&D, 1902-1980*. Cambridge University Press.

Hughes, T. P. (2004, June). *American Genesis: A Century of Invention and Technological Enthusiasm, 1870-1970*. University of Chicago Press.

Hunt, B. J. (Ed.). (2021). To Rule the Waves: Britain's Cable Empire and the Making of "Maxwell's Equations". In *Imperial Science: Cable Telegraphy and Electrical Physics in the Victorian British Empire* (pp. 216–271). Cambridge University Press.

Iaria, A., Schwarz, C., & Waldinger, F. (2018). Frontier knowledge and scientific production: Evidence from the collapse of international science. *The Quarterly Journal of Economics*, *133*(2), 927–991.

Jenkins, R., & Chandler, A. D. (1975). *Images and Enterprise: Technology and the American Photographic Industry, 1839 to 1925*. Johns Hopkins University Press.

Kandel, E., Kosenko, K., Morck, R., & Yafeh, Y. (2019). The Great Pyramids of America: A Revised History of US Business Groups, Corporate Ownership and Regulation, 1926-1950. *Strategic Management Journal*, *40*(5), 781–808.

Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, *3*(3), 303–20.

Khanna, T. (2000). Business Groups and Social Welfare in Emerging Markets: Existing Evidence and Unanswered Questions. *European Economic Review*, *44*(4-6), 748–761.

Khanna, T., & Palepu, K. (2000). Is Group Affiliation Profitable in Emerging Markets? An Analysis of Diversified Indian Business Groups. *The Journal of Finance*, *55*(2), 867–891.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, *132*(2), 665–712.

Lamoreaux, N. R., & Sokoloff, K. L. (1999). Inventors, Firms, and the Market for Technology in the Late Nineteenth and Early Twentieth Centuries. *Learning by Doing in Markets, Firms, and Countries* (pp. 19–60). University of Chicago Press.

Langmuir, I. (1913). The Effect of Space Charge and Residual Gases on Thermionic Currents in High Vacuum. *Physical Review*, *2*(6), 450–486.

Lassman, T. C. (2003). Industrial Research Transformed: Edward Condon at the Westinghouse Electric and Manufacturing Company, 1935-1942. *Technology and Culture*, *44*(2), 306–339.

Lichtenberg, F. R. (1986). *Private Investment in R&D to Signal Ability to Perform Government Contracts*. NBER Working Papers: No. 1974.

Lipartito, K. (1989). *The Bell System and Regional Business: The Telephone in the South, 1877-1920*. Johns Hopkins University Press.

Maclaurin, W. R., & Harman, R. J. (1949). *Invention & Innovation in the Radio Industry*. Macmillian.

Marx, M., & Fuegi, A. (2020). Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles. *Strategic Management Journal*.

Moser, P., Parsa, S., & San, S. (2022). Immigration, Science, and Invention. Lessons from the Quota Acts. *Available at SSRN 3558718*.

Moser, P., Voena, A., & Waldinger, F. (2014). German Jewish émigrés and US invention. *American Economic Review*, *104*(10), 3222–55.

Mowery, D., & Rosenberg, N. (1998). *Paths of innovation : Technological change in 20th-century America*. Cambridge ; New York : Cambridge University Press, 1998.

Mowery, D. C. (1983). The Relationship between Intrafirm and Contractual Forms of Industrial Research in American Manufacturing, 1900–1940. *Explorations in Economic History*, *20*(4), 351–374.

Mowery, D. C. (2009). Plus ca Change: Industrial R&D in the "Third Industrial Revolution". *Industrial and Corporate Change*, *18*(1), 1–50.

Mowery, D. C., & Simcoe, T. (2002). Is the Internet a US invention?—an economic and technological history of computer networking. *Research Policy*, *31*(8), 1369–1387.

Mueller, M. (1997). *Universal Service: Competition, Interconnection, and Monopoly in the Making of the American Telephone System*. MIT Press.

National Research Council. (1931). *Industrial Research Laboratories of the United States, Including Consulting Research Laboratories*. RR Bowker Company.

National Research Council. (1938). *Research: A National Resource : II. Industrial Research*. The National Academies Press.

Nicholas, T. (2010). The Role of Independent Invention in US Technological Development, 1880–1930. *The Journal of Economic History*, 57–82.

Nicholas, T. (2011). Did R&D Firms Used to Patent? Evidence from the First Innovation Surveys. *The Journal of Economic History*, *71*(4), 1032–1059.

Nutter, G. W., & Einhorn, H. A. (1969). *Enterprise Monopoly in the United States: 1899-1958*. Columbia University Press.

Perazich, G., & Field, P. M. (1940). *Industrial Research and Changing Technology*. WPA, Nat. Research Project.

Petralia, S., Balland, P.-A., & Rigby, D. L. (2016). Unveiling the Geography of Historical Patents in the United States from 1836 to 1975. *Scientific Data*, *3*(1), 1–14.

Reich, L. S. (1983). Irving Langmuir and the Pursuit of Science and Technology in the Corporate Environment. *Technology and Culture*, *24*(2), 199–221.

Reich, L. S. (1985). *The Making of American Industrial Research : Science and Business at GE and Bell, 1876-1926*. Cambridge University Press.

Reich, L. S. (1992). Lighting the Path to Profit: GE's Control of the Electric Lamp Industry, 1892-1941. *The Business History Review*, *66*(2), 305–334.

Rosenberg, N. (1990). Why Do Firms Do Basic Research (with Their Own Money)? *Research Policy*, *19*(2), 165–174.

Servos, J. W. (1994). Changing Partners: The Mellon Institute, Private Industry, and the Federal Patron. *Technology and Culture*, *35*(2), 221–257.

Soddy, F. (1907). The Wehnelt Kathode in a High Vacuum. *Nature*, *77*(1986), 53–54.

Stern, S. (2004). Do Scientists Pay to Be Scientists? *Management Science*, *50*(6), 835–853.

Stigler, G. J. (1949). Competition in the United States. *Five Lectures on Economic Problems* (pp.46–62). Longmans, Green and Co.

Terman, F. E. (1976). A Brief History of Electrical Engineering Education. *Proceedings of the IEEE*, *64*(9), 1399–1407.

Wilcox, C. (1940). Competition and Monopoly in American Industry. *Investigation of Concentration of Economic Power, T.N.E.C, Monograph No. 21*.

Wise, G. (1985). *Willis R. Whitney, General Electric, and the Origins of US industrial research*. Columbia University Press.

Wolfe, R. (2022, October 4). *Businesses Spent Over a Half Trillion Dollars for R&D Performance in the United States During 2020, a 9.1% Increase Over 2019.* (NSF 22-343). National Science Foundation. Alexandria, VA.

For Online Publication
Appendix A Data Appendix

Table A1: VARIABLE DEFINITION TABLE

Variable Name	Variable Definition	Source
Firm Performance		
ln(Market Value)	Log of market capitalization	Center for Research in Security Prices (CRSP)
Patents Within Top 5% Value (KPSS)	Number of annual focal firm patents within the top 5% of patent value	Kogan, Papanikolaou, Seru, and Stoffman (2017)
Patents Within Top 5% Forward Cites	Number of annual focal firm patents within the top 5% of forward patent citations	Google Patents
Corporate Investment in Science (r_0)		
Publication Count	Number of annual peer-reviewed scientific publications matched to focal firm	Microsoft Academic Graph (MAG)
Lab Employees	Number of annual laboratory employees matched to focal firm	Industrial Research Laboratories of the United States (1927,1931,1933,1938,1940)
AMS Scientists	Number of prominent scientists that are affiliated with focal firm	American Men of Science (1921)
Returns from Innovation (k)		
Dummy for Patent Cite to Science	Equals 1 for a firm whose patents make at least one citation to scientific publications during sample period, zero otherwise	Marx and Fuegi (2020)
Dummy for First Patent in CPC	Equals 1 for a firm that is issued the first patent in a CPC for a given year	Google Patents
Patent Importance (KPST)	Average "importance" of annual focal firm patents for given firm-year, where importance is measured by dividing the 10-year forward textual similarity by 5-year backward textual similarity of patents	Kelly, Papanikolaou, Seru, and Taddy (2021)
Market Share	Firm level average of focal firm annual sales normalized by annual sales in focal firm's 3-digit industry between 1926 and 1930	Kandel, Kosenko, Morck, and Yafeh (2019)
Dummy for Competitive Market	Dummy equalling 1 if focal firm's industry is classified as competitive and zero otherwise	Wilcox (1940)
Availability of Public Science (u)		
University Gap	Number of European scientific publications relevant to a firm divided by all (European and American) publications relevant to a firm (calculated for papers published between 1900 and 1920)	Clarivate Web of Science
University Gap (Cites)	Number of American journal citations to European journals divided by total American journal citations (calculated for citations made between 1900 and 1920)	Clarivate Web of Science
Control Variables		
ln(Assets)	Log of total assets for firm-year	Kandel, Kosenko, Morck, and Yafeh (2019) & Graham, Leary, and Roberts (2015)
Distance to Universities	Average distance of a firm to American universities granting graduate degrees in the sciences in 1930	Wilson (1932)

Table A2: Summary Statistics of Main Variables

	Obs	Mean	Median	Std Dev	Min	Max
Dummy for Patent Cite to Science	6990	0.07	0.00	0.25	0.00	1.00
Dummy for First Patent in CPC	6990	0.11	0.00	0.32	0.00	1.00
Patent Importance	4030	0.05	0.01	0.18	-0.48	1.31
Market Share	5955	0.05	0.00	0.12	0.00	1.00
Dummy for Competitive Market	6870	0.42	0.00	0.49	0.00	1.00
University Gap	6990	0.70	0.70	0.03	0.58	0.79
University Gap (Cites)	6990	0.48	0.48	0.08	0.28	0.78
Distance to Universities	6975	699.39	597.47	383.86	511.30	4456.13
Lab Employees	2310	45.00	0.00	221.55	0.00	4669.00
Publications Authored Per Year	6990	0.47	0.00	3.73	0.00	88.00
AMS Scientists	466	1.55	0.00	7.01	0.00	89.00
Patents Granted Per Year	6990	13.21	1.00	55.04	0.00	838.00
Patents Within Top 5% Value (KPSS)	3840	0.89	0.00	6.36	0.00	127.00
Patents Within Top 5% Forward Cites	6990	0.72	0.00	3.32	0.00	56.00
Total Assets ($MM)	4282	1373.64	413.79	3320.42	7.43	60114.66
Gross Income ($MM)	3242	864.31	272.59	1825.53	0.25	20655.93
Market Capitalization ($MM)	3840	1096.07	246.42	2905.64	0.69	37352.08

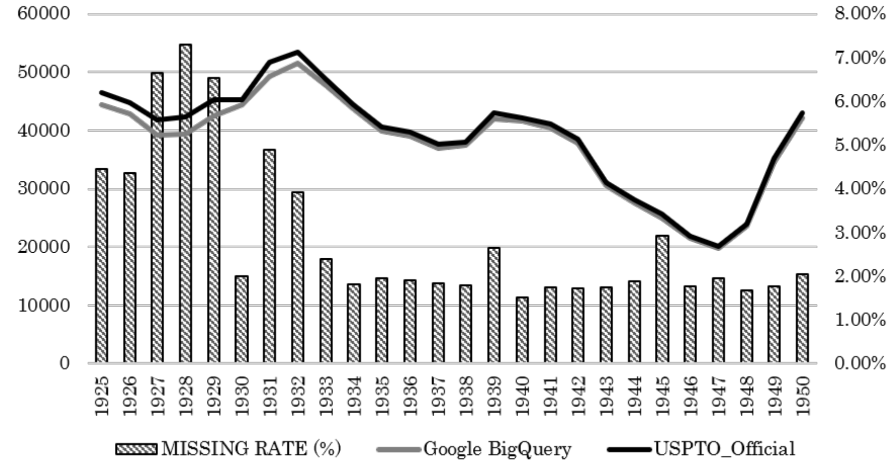
Notes: Observations are at the firm-year level. The sample period is between 1926 and 1940. All dollar amounts are deflated to 2005 dollars using https://www.measuringworth.com/datasets/usgdp12/result.php. Details on variable definition and data sources can be found in Appendix Table A1.

A.1 Matching Corporations to Patents, Publications, Laboratories and Scientists

A.1.1 Matching Corporations to Patents

Our patent data is sourced from the Google Patents dataset via Google BigQuery. We cross-check the number of utility patents granted each year with the official USPTO statistics for our sample period in Figure A1 to ensure that our data source does not have coverage issues.[48] We find that the missing rate is around 3.43%; there are an average of 42,476 utility patents granted per year between 1926 and 1940.

Figure A1: NUMBER OF PUBLISHED UTILITY PATENTS, 1925-1950



Source: The bar graph (right axis) plots the missing rate, defined as the difference in annual patent numbers between the USPTO official statistics and the Utility Patent (inventions) Column in the following source: https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm.

We extract the assignee field of the patents and standardize the names. We remove common prefixes and suffixes, such as “The,” “LLC,” “INC,” “A CORP OF”. We also standardize names common in certain industries such as petroleum (sometimes abbreviated as “petr”), utilities (“power” abbreviated as “pwr”), rail (“railway,” “railroad,” “rail” used interchangeably and variously abbreviated as “RC,” “RW,” “RD,” and “RC”) as well as more common names, such as “manufacturing” (“MFG”), “National” (“Nat’l Steel Corp.”), “American” (“Radio Corp of Amer”) and state abbreviations. The last standardization is important for our sample period because companies then were more often named after the states they operated

---

[48] USPTO official statistics for this period come from https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm.

in (for instance, "Delaware Lackawanna & Western Coal Co." or the "Pennsylvania Electric Company"). Furthermore, we find alternative names specific to certain firms such as the Standard Oil Company of Indiana (STANOLIND) and lab names for large companies such as AT&T's Bell Laboratories. Common abbreviations, such as RCA (Radio Corporation of America) and GE (General Electric), are also included. We then use a fuzzy string matching algorithm that calculates a length-adjusted Levenshtein distance. Using a fuzzy string matching algorithm is critical for patents from this period, as assignee names were not input electronically and are parsed through OCR.[49] Moreover, we manually check the names of 620 patentees with above 100 patents to include any matches that the string matching algorithm may still have missed.

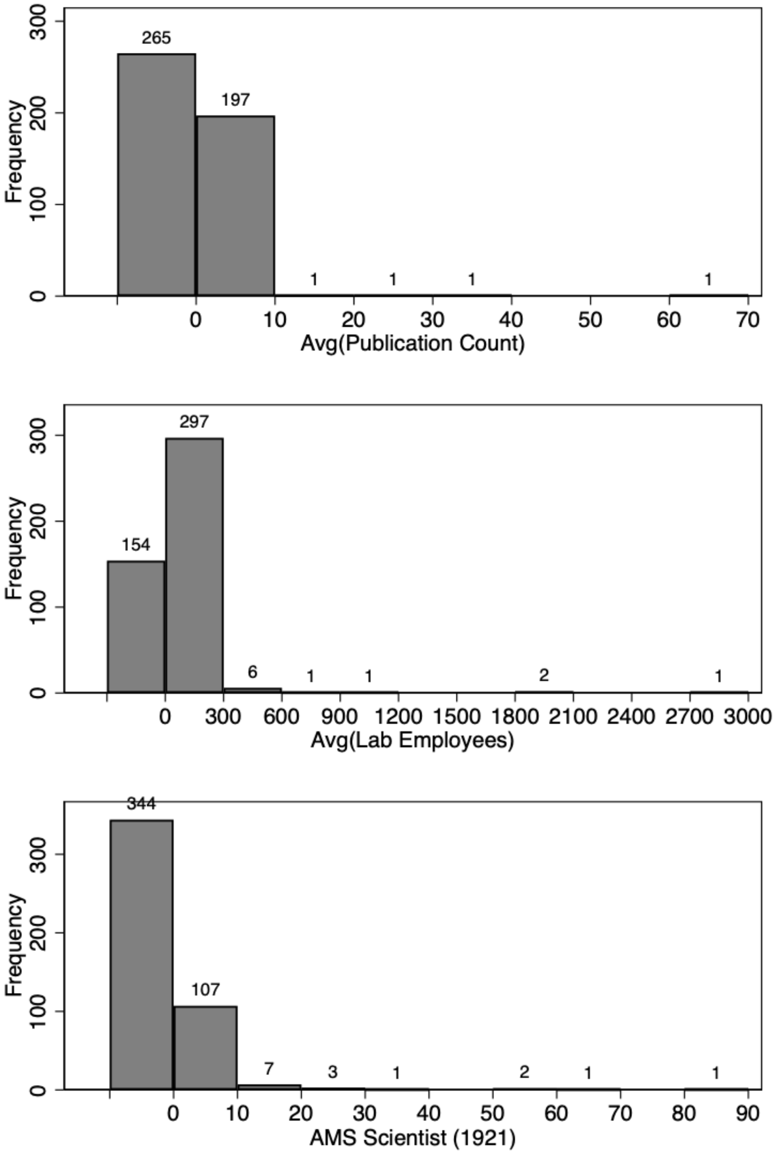
We match 318 firms found in the B&M sample to 64,523 patents. We also add 2,344 additional patents matched to 38 CRSP firms that were not matched in Kogan, Papanikolaou, Seru, and Stoffman (2017).[50]

---

[49] As an example, the SOCONY Vacuum Oil Company is "misspelled" in the Google Patent data as: SCONY VACUUM OIL CO INC, SOCCNY VACUUM OIL CO INC, SOCENY VACUUM OIL CO IN, SOCONEY VACUUM OIL CO INC, SOCONY VACUNM OIL CO INC, SOCONY VAEUUM OIL CO INC, SOCONYVACUUM OIL CO INC, SOECNY VACUUM OIL CO INC, SOEONY VACUUM OIL CO INC, and SONCONY VACUUM OIL CO INC. The fuzzy string matching algorithm is still able to recover these matches.

[50] Kogan, Papanikolaou, Seru, and Stoffman (2017) match 60,493 patents to 368 CRPS firms between 1926 and 1940, which we also add to our sample.

Figure A2: HETEROGENEITY OF CORPORATE SCIENCE



Notes: The upper histogram bins the number of publications authored by firms in our sample. 265 firms (the leftmost bar) do not author any scientific publications between 1926 and 1940. The middle histogram bins the number of personnel employed at corporate laboratories for firms in our sample. 154 firms (the leftmost bar) report no employed lab personnel in our sample period. The lower histogram bins the number of scientists in AMS affiliated with firms in our sample. 344 firms (the leftmost bar) do not employ any AMS scientists in 1921.

A.1.2 Matching Corporations to Publications

Table A3: AMERICAN CORPORATE PUBLICATIONS (TOP 20)

Firm Name	Paper Count
GENERAL ELECTRIC CO	1146
AMERICAN TELEPHONE & TELEG CO	658
WESTINGHOUSE ELECTRIC & MFG CO	466
RADIO CORP AMER	207
EASTMAN KODAK CO	173
SQUIBB E R & SONS	112
WESTERN ELECTRIC COMPANY, INC	100
COMMONWEALTH EDISON CO	49
SWIFT & CO	44
HUMBLE OIL AND REFINING COMPANY	41
PROCTER & GAMBLE CO	40
SHARP & DOHME INC	40
PARKE DAVIS & CO	37
SOUTHERN CALIFORNIA EDISON CO	34
PHILADELPHIA ELECTRIC CO	33
CORNING GLASS WORKS	30
WESTERN UNION TELEGRAPH CO	29
WESTINGHOUSE LAMP COMPANY	29
DETROIT EDISON CO	29
PHILADELPHIA ELECTRIC POWER CO	29

Notes: The table presents the number of scientific publications in MAG between 1900 and 1940 matched to our sample firms. The top 20 publishing firms are included.

Our publication data is sourced from Microsoft Academic Graph. We first download all author affiliations for papers published between 1900 and 1940.[51] We run the same fuzzy string matching algorithm as above and manually check matches above a threshold score. Unlike patents, corporate publications are also often published under the name of the lab, which may not always correspond to the name of the firm. Therefore, we add names of prominent corporate laboratories such as Bell Labs and the Edgar C Bain Lab (for U.S. Steel) as name variants. To prevent false positive matches, we check that charitable organizations and university labs are not mismatched to the company. For instance, a 1934 publication by the "Eastman Laboratory of Physics" has high textual similarity to Eastman Kodak, but is actually part of the Massachusetts Institute of Technology, with no ties to the

---

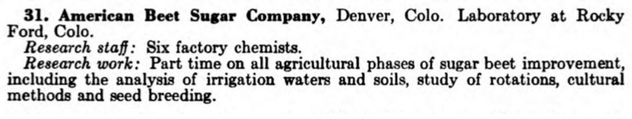
[51] Though our main sample runs from 1926 to 1940, publications data before 1926 are used in the analyses in Section 2.

firm. We also cross-tabulated the publication field of the company with its industry as a sanity check: we confirm, for instance, the wholesale and retail industry has scientific publications because the Boots Pure Drug Company (classified under this industry) published 29 articles ranging from the chemical sciences to clinical medicine.

A.1.3 Matching Corporations to Industrial Research Laboratories

We download the PDF files for the 1927, 1931, 1933 and 1938 editions of the NRC's Industrial Research Laboratory directory from Hathitrust. Since lab entries in the directory are of varying length (e.g., a stub for the American Beet Sugar Company (figure A3) vs 2 pages for DuPont (figure A4)) and the fields are not sorted into metadata, the use of automated string matching algorithms is inefficient. However, since the entries are listed alphabetically, the directories are still amenable to manual matching. We enlisted two research assistants that manually searched through the directory to gather the name of the lab and the number of personnel employed at them. Though the directory also lists the type of personnel employed (e.g., chemists, physicists, etc.), these are not standardized by training or salary level, making it difficult to compare across firms. Therefore, we only use the total number of personnel as the indicator of investment in science for the analysis.

Figure A3: 1933 IRL ENTRY FOR AMERICAN BEET & SUGAR COMPANY

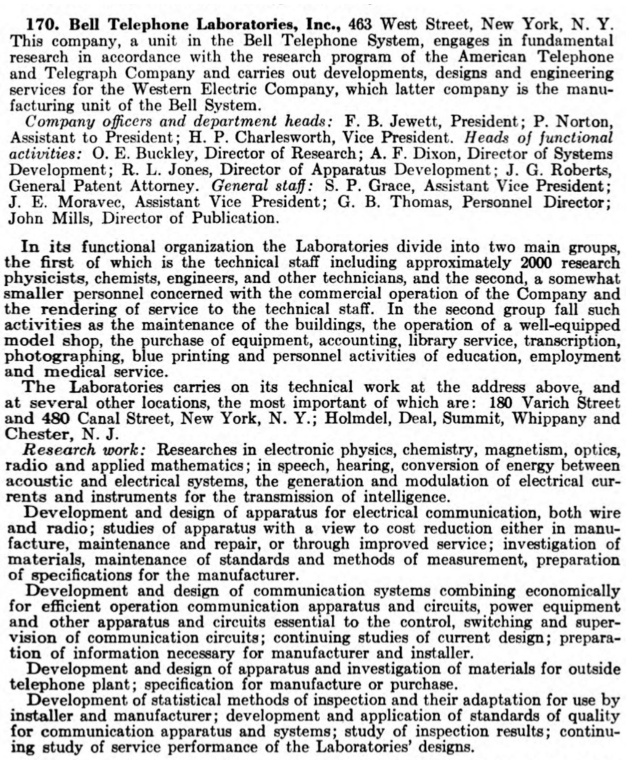


31. American Beet Sugar Company, Denver, Colo. Laboratory at Rocky Ford, Colo.

Research staff: Six factory chemists.

Research work: Part time on all agricultural phases of sugar beet improvement, including the analysis of irrigation waters and soils, study of rotations, cultural methods and seed breeding.

Figure A4: 1933 IRL ENTRY FOR AT&T BELL LABS



170. Bell Telephone Laboratories, Inc., 463 West Street, New York, N. Y. This company, a unit in the Bell Telephone System, engages in fundamental research in accordance with the research program of the American Telephone and Telegraph Company and carries out developments, designs and engineering services for the Western Electric Company, which latter company is the manufacturing unit of the Bell System.

Company officers and department heads: F. B. Jewett, President; P. Norton, Assistant to President; H. P. Charlesworth, Vice President. *Heads of functional activities:* O. E. Buckley, Director of Research; A. F. Dixon, Director of Systems Development; R. L. Jones, Director of Apparatus Development; J. G. Roberts, General Patent Attorney. *General staff:* S. P. Grace, Assistant Vice President; J. E. Moravec, Assistant Vice President; G. B. Thomas, Personnel Director; John Mills, Director of Publication.

In its functional organization the Laboratories divide into two main groups, the first of which is the technical staff including approximately 2000 research physicists, chemists, engineers, and other technicians, and the second, a somewhat smaller personnel concerned with the commercial operation of the Company and the rendering of service to the technical staff. In the second group fall such activities as the maintenance of the buildings, the operation of a well-equipped model shop, the purchase of equipment, accounting, library service, transcription, photographing, blue printing and personnel activities of education, employment and medical service.

The Laboratories carries on its technical work at the address above, and at several other locations, the most important of which are: 180 Varich Street and 480 Canal Street, New York, N. Y.; Holmdel, Deal, Summit, Whippany and Chester, N. J.

Research work: Researches in electronic physics, chemistry, magnetism, optics, radio and applied mathematics; in speech, hearing, conversion of energy between acoustic and electrical systems, the generation and modulation of electrical currents and instruments for the transmission of intelligence.

Development and design of apparatus for electrical communication, both wire and radio; studies of apparatus with a view to cost reduction either in manufacture, maintenance and repair, or through improved service; investigation of materials, maintenance of standards and methods of measurement, preparation of specifications for the manufacturer.

Development and design of communication systems combining economically for efficient operation communication apparatus and circuits, power equipment and other apparatus and circuits essential to the control, switching and supervision of communication circuits; continuing studies of current design; preparation of information necessary for manufacturer and installer.

Development and design of apparatus and investigation of materials for outside telephone plant; specification for manufacture or purchase.

Development of statistical methods of inspection and their adaptation for use by installer and manufacturer; development and application of standards of quality for communication apparatus and systems; study of inspection results; continuing study of service performance of the Laboratories' designs.

Table A4: CORPORATE SCIENTIFIC PUBLICATIONS, BY OECD SUBFIELD

OECD Subfield	Number of Firms	Number of Papers	Average Forward Publication Citations
1.03 Physical sciences and astronomy	39	433	1.73
1.06 Biological sciences	27	71	1.52
2.03 Mechanical engineering	57	148	1.37
2.02 Electrical eng, electronic eng	63	1268	0.91
1.04 Chemical sciences	59	267	0.88
4.01 Agriculture, forestry, fisheries	5	8	0.77
1.02 Computer and information sciences	29	77	0.65
1.01 Mathematics	21	77	0.61
1.07 Other natural sciences	2	2	0.60
2.05 Materials engineering	40	152	0.59
3.02 Clinical medicine	32	112	0.55
2.08 Environmental biotechnology	6	6	0.51
1.05 Earth and related environmental sciences	34	96	0.44
2.11 Other engineering and technologies	44	97	0.37
2.06 Medical engineering	4	21	0.31
3.01 Basic medical research	19	23	0.26
3.03 Health sciences	12	19	0.22
2.01 Civil engineering	30	73	0.19
2.07 Environmental engineering	50	169	0.19
2.04 Chemical engineering	19	22	0.13
4.02 Animal and dairy science	7	10	0.11
4.03 Veterinary science	2	3	0.04
4.05 Other agricultural science	8	10	0.03
Not Available	37	99	0.02

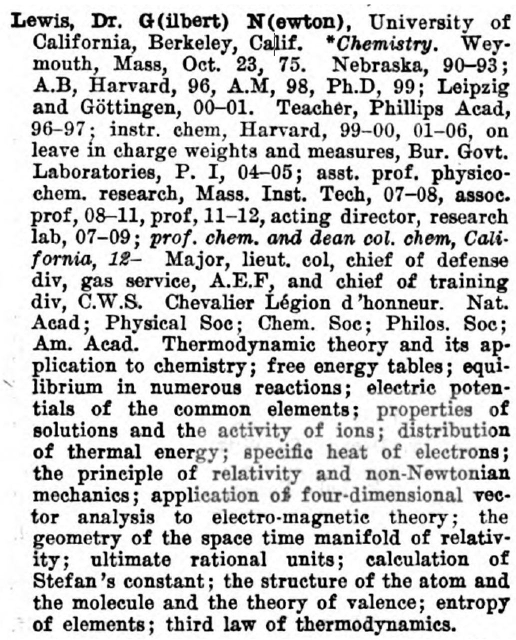
Notes: Observations are at OECD subfield level for years between 1926 and 1940. "Number of Firms" counts the number of firms publishing at least one article in the focal field. "Number of Papers" counts the number of total papers in the focal field. "Average Forward Publication Cites" take the field-level average of the normalized forward citations. Forward citations are normalized by the average number of forward citations received by all publications published in the focal publication's year.

A.1.4 Matching Corporations to American Men of Science Directory

The AMS directory lists information on each scientist in a consistent manner: the last name is followed by the title, first name, current employment and residence and main discipline. Information on date and place of birth, alma mater, past employment and membership in professional societies follow. The final item in each entry is a detailed list of keywords that describe the focal scientist's research interests. We make use of a dataset that manually inputs this information into spreadsheet format, which has been used in recent works such as Moser and Kim (2022), Moser and Parsa (2022), and Moser et al. (2022). A total of 9,557 scientists are present and up to 30 employers per scientist are identified from the 1921 edition.

We manually search for the names of the 466 firms in our sample to identify when a scientist has been employed.

Figure A5: American Men of Science Entry for Gilbert Lewis (1921)



Lewis, Dr. G(ilbert) N(ewton), University of California, Berkeley, Calif. **Chemistry.* Weymouth, Mass, Oct. 23, 75. Nebraska, 90–93; A.B, Harvard, 96, A.M, 98, Ph.D, 99; Leipzig and Göttingen, 00–01. Teacher, Phillips Acad, 96–97; instr. chem, Harvard, 99–00, 01–06, on leave in charge weights and measures, Bur. Govt. Laboratories, P. I, 04–05; asst. prof. physico-chem. research, Mass. Inst. Tech, 07–08, assoc. prof, 08–11, prof, 11–12, acting director, research lab, 07–09; *prof. chem. and dean col. chem, California, 12–* Major, lieut. col, chief of defense div, gas service, A.E.F, and chief of training div, C.W.S. Chevalier Légion d'honneur. Nat. Acad; Physical Soc; Chem. Soc; Philos. Soc; Am. Acad. Thermodynamic theory and its application to chemistry; free energy tables; equilibrium in numerous reactions; electric potentials of the common elements; properties of solutions and the activity of ions; distribution of thermal energy; specific heat of electrons; the principle of relativity and non-Newtonian mechanics; application of four-dimensional vector analysis to electro-magnetic theory; the geometry of the space time manifold of relativity; ultimate rational units; calculation of Stefan's constant; the structure of the atom and the molecule and the theory of valence; entropy of elements; third law of thermodynamics.

Source: Entry on Gilbert Lewis from the 1921 edition of the American Men of Science Directory.

A.2 Measuring Geographic Distances Between Firms and Universities

A.2.1 Geolocation Data on Firms

We collect the addresses of B&M firms and their subsidiaries from the 1929-1930 Moody's Public Utilities, Railroad and Industrial Manuals. Each entry has a section on the firm's management team ("Officers"), in which the firm's office location is indicated. In the case of firms with multiple offices, we use the main office as the firm's location (in the case of the Porto Rico Telephone Company in Figure A7, the location is San Juan, Puerto Rico).

For firms that are not included in the B&M sample, we use the patents matched to their assignee names by their patent numbers to the HistPat dataset (Petralia et al., 2016). HistPat is a publicly available dataset that collects geolocational data on patent inventors and assignees for U.S. patents prior to 1976 through a text mining algorithm. We use version 8.0 of the dataset[52] and extract the FIPS (Federal Infor-

---

[52] Available from https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/

mation Processing Standards) County Codes of the assignee of the firms' patents. Where there are multiple counties associated with a firm's patents, we choose the county that appears most frequently.

A.2.2 Geolocation Data on Universities

We collect the location of all universities in the United States that were granting doctorates in the natural sciences from Wilson (1932).[53] While select universities were publishing catalogs of theses, a nation-wide catalog did not begin until 1912, when the Librarian of Congress began compiling doctoral dissertations from all degree-granting institutions. The Library circulated letters to "all universities listed in the latest "Report to the Commissioner of education" as maintaining graduate departments." to receive "every thesis printed," with the aim of acquiring, classifying and cataloguing them (Flagg, 1913, p.7). The catalog is prepared from this annual list, complete with subject headings for each dissertation. We use the 1932 volume, which contains dissertations that were submitted between January 1931 and September 1932. Based on subject headings of the dissertations, we removed universities that do not grant doctorates in the natural sciences such as the "Peabody College for Teachers" and the "Dropsie College for Hebrew and Cognate Learning". At the end of the process, we find 41 universities granting doctorates in the natural sciences in 1930 and manually collect the addresses of the institutions from the web, under the assumption that university locations have not changed over time.

A.2.3 Measuring Distances

For consistency with the HistPat database, we gather the FIPS County codes for offices addresses from Moody's (for the B&M firms) and for university addresses from Flagg (1913). We then calculate the latitude and longitude of the centroid of all FIPS and calculate the geodesic distances between all 466 firm and 41 university combinations using the Stata module "geodist".[54]

A.3 Corporate Ownership Data

We define a business group as a collection of three or more listed firms under common ownership. For years in our sample where ownership data are not collected (i.e., years other than 1926,29,32,37,40), we impute business group affiliation years if they do not change between consecutive collection years.[55] In addition to using business group data, we measure levels of diversification by calculating Herfindahl (HHI) indices of sales distributions across 3-digit SIC industries at the ultimate owner level.[56] For each ultimate owner-year, we calculate the share of each 3-

---

BPC15W

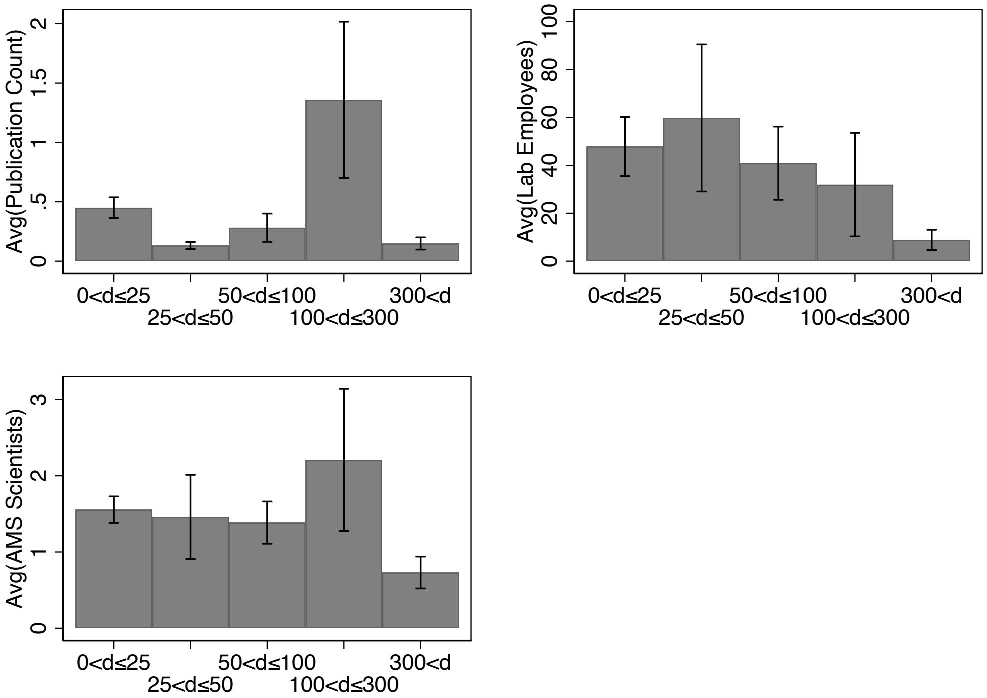
[53] Available form https://babel.hathitrust.org/cgi/pt?id=uc1.b3509036&view=1up&seq=7).

[54] https://ideas.repec.org/c/boc/bocode/s457147.html

[55] That is, if a firm is controlled by General Electric in 1926 and 1929, years 1927 and 1928 are imputed for the firm as GE affiliate years.

[56] The ultimate owner firm is the "apex firm" at the end of a control chain.

Figure A6: Corporate Publications, Lab Personnel and Scientists, by Distance from Universities



Notes: This bar graph plots the mean and 95% confidence intervals of scientific publications, lab employees and men of science (from AMS 1921) for our sample firms, categorized by their proximity to American universities producing doctorates in the natural sciences. *d* refers to the minimum geodesic distance in miles between a firm and universities. For instance, the "0<d≤25" group refers to firms that have at least one university within 25 miles.

digit industry out of total sales and sum the squared shares across industries. A Herfindahl index of 1 implies that the group of firms owned by the ultimate owner derives all of its sales from a single industry.[57]

A.3.1 Control Chains

We use Moody's Manuals to track companies controlling, or controlled by, the 200 companies on the B&M list. In each volume, a company report is followed by reports on its controlled subsidiaries (which are identified without an explicitly specified control threshold held by the controlling company). For example, if company A controls company B and company B, in turn, controls company C, and all three firms belong to the railroad sector, the A-B-C control chain will appear in Moody's Railroads Manual in the same sequence with the identity of the corporate controller usually reported next to the company name. We examine if one or more companies are controlled by another corporation included in the original list and, if this is the case, combine their control chains. Therefore, each control chain in our sample is a long sequence of firms consisting of an apex corporation and its subsidiaries, each of which has control over the next one. In most cases, control chains include firms belonging to the same industrial category (e.g., railroads), but there are occasionally multiple control chains in different categories with the same ultimate owner as well (e.g., a few cases of public utility apex companies controlling industrial companies).

A.3.2 Ultimate Controlling Shareholders

Moody's Manuals do not provide any information on the identity of the controllers of apex firms. To identify the owners of apex corporations that are not controlled by any other entity, we use the following sources:

1. For the 1926-1929 period: Pinchot (1928), the Wall Street Journal (WSJ) and the New York Times (NYT) archives, as well as additional sources, such as internet searches, historical documents, corporate files, www.archives.org and www.fundinguniverse.com.

2. For the 1929-1932 period: Table XII, Berle and Means (1932), Bonbright and Means (1932), Buchanan (1936), Lundberg (1937), the Encyclopedia of American Business History (2006), the WSJ and NYT archives and www.fundinguniverse.com.

3. For the 1937-1940 period: National Resources Committee (1939, Chapter IX and Appendix 13) and TNEC (1940).

---

[57] It is important to note that we calculate HHI for any multi-firm entity regardless of the number of listed affiliates. Therefore, the calculation of HHI is not restricted to business groups only, as defined in Kandel et al. (2019)

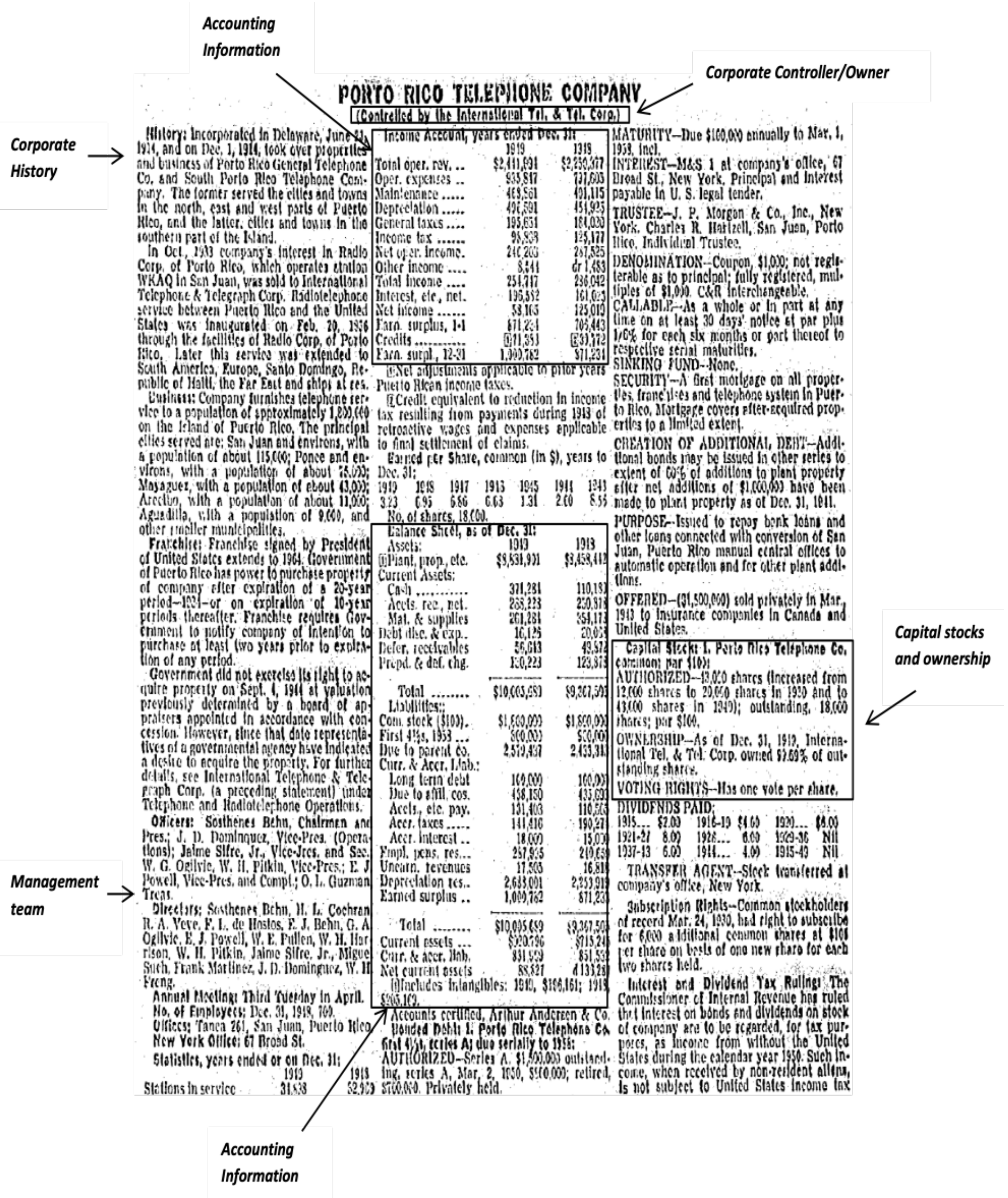
A.3.3 Corporate Historical Documents and Data Sources

- Bureau of Economic Analysis (BEA, 1958), U.S. Department of Commerce, Benchmark Federal Trade Commission (FTC) Annual Reports: www.ftc.gov/os/annualreports/index.shtm
- Input-Output Data: Historical SIC Data, www.bea.gov/industry/io_histsic.htm
- Interstate Commerce Commission (ICC) Reports
- Moody's Manuals, 1926-1940: http://webreports.mergent.com/
- Statistics of Income: http://www.irs.gov/pub/irs-soi/
- National Association of Railroad and Utility Commissioners
- National Resources Committee (NRC) (1939), The Structure of the American Economy (Washington, DC: U.S. Government Print Office)
- Regulation of Stock Ownership in Railroads, 71st Congress, 3d Session, House Report No. 2789, Vol.2, February 1931
- Securities and Exchange Commission (SEC) Annual Reports: www.sec.gov/about/annrep.shtml
- Survey of American Listed Corporations: Reported Information on Registrants with the SEC under the Securities Exchange Act of 1934, 1939-40
- Temporary National Economic Committee (TNEC), (1940), The Distribution of Ownership in the 200 Largest Nonfinancial Corporations, monograph 29 (1-2) (Washington, DC: U.S. Government Printing Office): http://www.bpl.org/govinfo/online-collections/federal-executive-branch/temporary-national-economic-committee-1938-1941/
- Twentieth Century Fund, Committee on Taxation (1937), Facing the Tax Problem (New York: Twentieth Century Fund)

A.3.4 Corporate Histories

- http://www.Archive.org
- Encyclopedia of American Business History (Facts on File, 2005): http://www.Fundinguniverse.com
- The New York Times Archives: http://www.nytimes.com/ref/membercenter/nytarchive.html

Figure A7: A Moody's Manuals entry: The Porto Rico Telephone Company, 1949



PORTO RICO TELEPHONE COMPANY

(Controlled by the International Tel. & Tel. Corp.)

History: Incorporated in Delaware, June 21, 1914, and on Dec. 1, 1914, took over properties and business of Porto Rico General Telephone Co. and South Porto Rico Telephone Company. The former served the cities and towns in the north, east and west parts of Puerto Rico, and the latter, cities and towns in the southern part of the Island.

In Oct., 1933 company's interest in Radio Corp. of Porto Rico, which operates station WKAQ in San Juan, was sold to International Telephone & Telegraph Corp. Radiotelephone service between Puerto Rico and the United States was inaugurated on Feb. 20, 1936 through the facilities of Radio Corp. of Porto Rico. Later this service was extended to South America, Europe, Santo Domingo, Republic of Haiti, the Far East and ships at sea.

Business: Company furnishes telephone service to a population of approximately 1,800,000 on the Island of Puerto Rico. The principal cities served are: San Juan and environs, with a population of about 115,000; Ponce and environs, with a population of about 75,000; Mayaguez, with a population of about 43,000; Arecibo, with a population of about 11,000; Aguadilla, with a population of 9,000, and other smaller municipalities.

Franchise: Franchise signed by President of United States extends to 1964. Government of Puerto Rico has power to purchase property of company after expiration of a 20-year period—1934—or on expiration of 10-year periods thereafter. Franchise requires Government to notify company of intention to purchase at least two years prior to expiration of any period.

Government did not exercise its right to acquire property on Sept. 4, 1944 at valuation previously determined by a board of appraisers appointed in accordance with concession. However, since that date representatives of a governmental agency have indicated a desire to acquire the property. For further details, see International Telephone & Telegraph Corp. (a preceding statement) under Telephone and Radiotelephone Operations.

Officers: Sosthenes Behn, Chairman and Pres.; J. D. Dominguez, Vice-Pres. (Operations); Jaime Sifre, Jr., Vice-Pres. and Sec.; W. G. Ogilvie, W. H. Pitkin, Vice-Pres.; E. J. Powell, Vice-Pres. and Compt.; O. L. Guzman, Treas.

Directors: Sosthenes Behn, H. L. Cochran, R. A. Veve, F. L. de Hostos, E. J. Behn, G. A. Ogilvie, E. J. Powell, W. E. Pullen, W. H. Harrison, W. H. Pitkin, Jaime Sifre, Jr., Miguel Such, Frank Martinez, J. D. Dominguez, W. H. Freng.

Annual Meeting: Third Tuesday in April.

No. of Employees: Dec. 31, 1948, 700.

Offices: Tanca 261, San Juan, Puerto Rico.

New York Office: 67 Broad St.

Statistics, years ended or on Dec. 31:

	1949	1948
Stations in service	31,838	32,909

Income Account, years ended Dec. 31:

	1949	1948
Total oper. rev.	$2,411,891	$2,250,377
Oper. expenses	935,817	737,603
Maintenance	463,561	401,115
Depreciation	496,591	454,925
General taxes	195,651	184,030
Income tax	95,808	125,177
Net oper. income	240,203	287,525
Other income	8,544	dr 1,483
Total income	248,747	286,042
Interest, etc., net	195,582	161,023
Net income	53,165	125,019
Earn. surplus, 1-1	871,231	705,443
Credits	[a]71,351	[b]39,772
Earn. surpl., 12-31	1,000,762	871,231

[a]Net adjustments applicable to prior years Puerto Rican income taxes.

[b]Credit equivalent to reduction in income tax resulting from payments during 1948 of retroactive wages and expenses applicable to final settlement of claims.

Earned per Share, common (in $), years to Dec. 31:

1949	1948	1947	1946	1945	1944	1943
3.23	6.95	6.86	6.63	1.31	2.00	8.55

No. of shares, 18,000.

Balance Sheet, as of Dec. 31:

Assets:	1949	1948
[c]Plant, prop., etc.	$9,531,901	$8,458,412
Current Assets:		
Cash	371,281	110,181
Accts. rec., net	268,223	250,878
Mat. & supplies	281,284	354,173
Debt disc. & exp.	16,125	20,053
Defer. receivables	56,613	49,574
Prepd. & def. chg.	130,223	123,873
Total	$10,655,650	$9,367,503
Liabilities:		
Com. stock ($100)	$1,800,000	$1,800,000
First 4½s, 1958	800,000	900,000
Due to parent co.	2,579,437	2,433,313
Curr. & Accr. Liab.:		
Long term debt	100,000	100,000
Due to affil. cos.	458,150	435,693
Accts., etc. pay.	131,403	110,563
Accr. taxes	141,416	190,271
Accr. interest	18,000	15,000
Empl. pens. res.	257,935	210,659
Unearn. revenues	17,505	16,818
Depreciation res.	2,683,001	2,253,919
Earned surplus	1,000,762	871,231
Total	$10,655,650	$9,367,503
Current assets	$920,796	$715,246
Curr. & accr. liab.	831,959	851,532
Net current assets	88,837	d133,286

[c]Includes intangibles: 1949, $196,161; 1948, $205,169.

Accounts certified, Arthur Andersen & Co.

Bonded Debt: 1. Porto Rico Telephone Co. first 4½s, series A; due serially to 1958:

AUTHORIZED—Series A, $1,500,000; outstanding, series A, Mar. 2, 1950, $800,000; retired, $700,000. Privately held.

MATURITY—Due $100,000 annually to Mar. 1, 1958, incl.

INTEREST—M&S 1 at company's office, 67 Broad St., New York. Principal and interest payable in U. S. legal tender.

TRUSTEE—J. P. Morgan & Co., Inc., New York. Charles R. Hartzell, San Juan, Porto Rico, Individual Trustee.

DENOMINATION—Coupon, $1,000; not registerable as to principal; fully registered, multiples of $1,000. C&R interchangeable.

CALLABLE—As a whole or in part at any time on at least 30 days' notice at par plus 1/6% for each six months or part thereof to respective serial maturities.

SINKING FUND—None.

SECURITY—A first mortgage on all properties, franchises and telephone system in Puerto Rico. Mortgage covers after-acquired properties to a limited extent.

CREATION OF ADDITIONAL DEBT—Additional bonds may be issued in other series to extent of 60% of additions to plant property after net additions of $1,000,000 have been made to plant property as of Dec. 31, 1941.

PURPOSE—Issued to repay bank loans and other loans connected with conversion of San Juan, Puerto Rico manual central offices to automatic operation and for other plant additions.

OFFERED—($1,500,000) sold privately in Mar., 1943 to insurance companies in Canada and United States.

Capital Stock: 1. Porto Rico Telephone Co. common; par $100:

AUTHORIZED—43,000 shares (increased from 12,000 shares to 20,000 shares in 1920 and to 43,000 shares in 1949); outstanding, 18,000 shares; par $100.

OWNERSHIP—As of Dec. 31, 1949, International Tel. & Tel. Corp. owned 99.69% of outstanding shares.

VOTING RIGHTS—Has one vote per share.

DIVIDENDS PAID:

1915...	$2.00	1916-19	$4.00	1920...	$8.00
1921-27	8.00	1928...	6.00	1929-36	Nil
1937-43	6.00	1944...	4.00	1945-49	Nil

TRANSFER AGENT—Stock transferred at company's office, New York.

Subscription Rights—Common stockholders of record Mar. 24, 1930, had right to subscribe for 6,000 additional common shares at $105 per share on basis of one new share for each two shares held.

Interest and Dividend Tax Ruling: The Commissioner of Internal Revenue has ruled that interest on bonds and dividends on stock of company are to be regarded, for tax purposes, as income from without the United States during the calendar year 1950. Such income, when received by non-resident aliens, is not subject to United States income tax

Notes: This figure reproduces the 1949 entry for the Porto Rico Telephone Company (http://webreports.mergent.com).

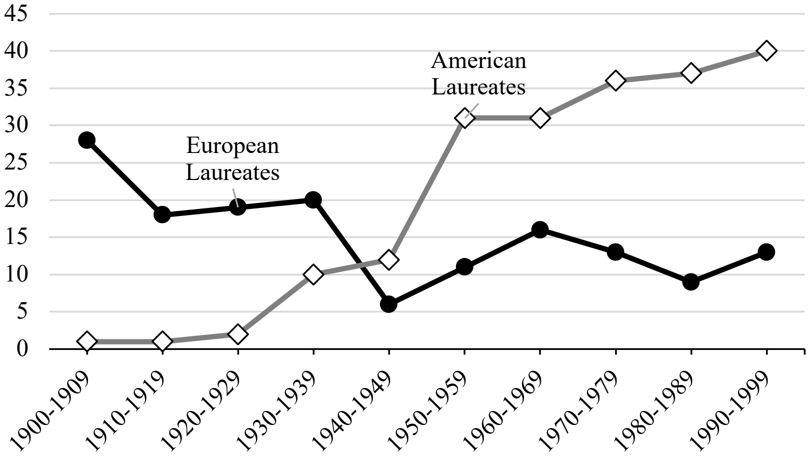
- The Wall Street Journal Archives: http://pqasb.pqarchiver.com/wsj/search.html

Appendix B Gaps in University Science

B.1 Details on Calculating Gaps in University Science

Figure B1 plots the number of Nobel Laureates in the natural sciences by their country of origin and year of award. European laureates outnumber American laureates until the onset of World War II. This is consistent with our assumption about the state of American Academia at the beginning of our sample period and motivates the comparisons to Europe to measure backwardness of American science in select subfields.

Figure B1: NUMBER OF NATURAL SCIENCE NOBEL PRIZE LAUREATES, BY CITIZENSHIP AT AWARD



Notes: The line graph plots the number of total laureates in the Nobel Prize for Physics, Chemistry, and Physiology/Medicine. The home countries of the winners are coded based on the classification by the Encyclopedia Britannica (please see https://www.britannica.com/topic/Winners-of-the-Nobel-Prize-for-Physics-1856942 for page for Physics). According to the source, "Nationality given is the citizenship of recipient at the time award was made."

Figure B2 provides an example of how the publications-based gap calculation is done for Black & Decker Company, which patents in four patent classes between 1926 and 1940. Half of its patents are in tools and bench devices (B25B) and the rest of its patents are equally divided among three patent classes (B23F , G01G and H02K). Of these, only Dynamo-electric devices (H02K) has patents that make Non-Patent Literature (NPL) citations to the scientific literature between 1947 and 1957. Among all NPL citations made from H02K during this period, 80% are made to Electrical Engineering and 20% are made to Materials Engineering. Hence, the

European papers relevant to the IPC are:

$$Eur.\ Papers_{H02K,EE} + Eur.\ Papers_{H02K,ME} = 0.8 \times 143 + 0.2 \times 15456 = 3205.6 \tag{B1}$$

while the American papers relevant to the IPC are $US\ Papers_{H02K,EE} + US\ Papers_{H02K,ME} = 0.8 \times 125 + 0.2 \times 3806 = 862$. The European papers relevant to the firm are then calculated as $Eur.\ Papers_{Black\&Decker,H02K} + Eur.\ Papers_{Black\&Decker,G01G} + Eur.\ Papers_{Black\&Decker,B23F} + Eur.\ Papers_{Black\&Decker,B25B} = .167 \times 3205.6 + .167 \times 0 + .167 \times 0 + .5 \times 0 = 538.5$.

The American papers relevant to the firm are similarly calculated as $US\ Papers_{Black\&Decker,H02K} + US\ Papers_{Black\&Decker,G01G} + US\ Papers_{Black\&Decker,B23F} + US\ Papers_{Black\&Decker,B25B} = .167 \times 862 + .167 \times 0 + .167 \times 0 + .5 \times 0 = 144$. It follows that *Gap in university science, 1900-20* value for Black & Decker is $538.5/(538.5 + 144) = .79$.

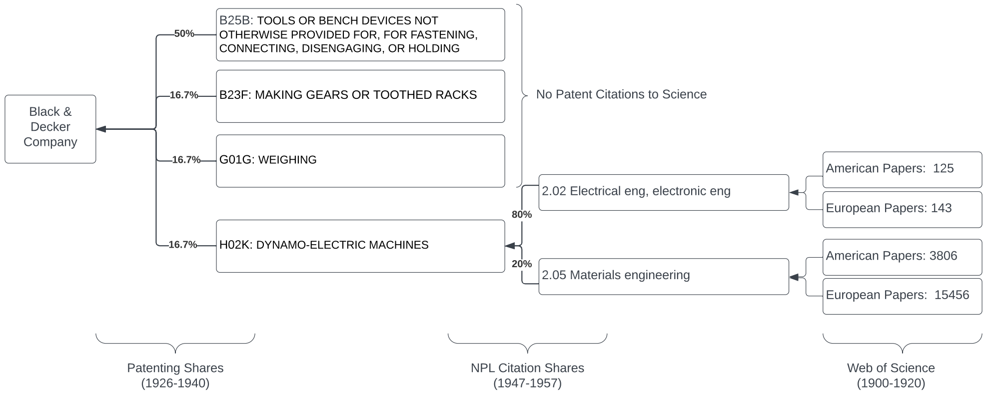
For the period between 1900 to 1920, Microsoft Academic Graph data do not record the country of publication. Also, we find that the affiliations sections rarely list the full address of the author for this period, which leads MAG to omit country data from affiliation data. We therefore rely on Clarivate Web of Science, which has previously been used for research on the impacts of World War I on scientific production (Iaria, Schwarz, & Waldinger, 2018). Of 307,847 publications listed in Web of Science, 15% (44,356) have country data. We code each country as American, European and Rest of the World. For the remaining 85% of publications without country information, we match the names of the authors to the 1906 and 1921 versions of the Cattell directory and classify those authors found in the directory as American (and the rest as European).

Another way to measure scientific gaps is by the number of citations made to publications in European journals by American journals. We classify 244 journals indexed in the WoS Science Citation Index - Expanded (SCI-EXPANDED) as "American" or "European" based on name and web searches. We first classify journals with non-English and non-Latin names (e.g., Zeitshcrift für Physik) as European. We also classify journals with the name "American" in it as American (e.g., the American Heart Journal). We then manually classify the remaining journals by web searches. Where a full history of the journal is available, we classify the journal's home country as the place where its publisher/publishing academic society is. For instance, "Bacteriological Reviews" is a journal that was published by the American Society of Microbiology.[58] When publisher information is not available, we use the nationality of the founding members to classify the journal. 230 journals out of the 244 are classified, 111 (45%) of which are American.

---

[58] https://en.wikipedia.org/wiki/Microbiology_and_Molecular_Biology_Reviews

Figure B2: Publication-Based Scientific Gap Calculation for Black & Decker



Notes: American and European paper numbers refer to papers published between 1900 and 1920 weighted by forward citations received up until 2019.

For articles published between 1900 and 1920, we count the number of citations made by “American” journals to “European” journals in the same period. This constitutes a measure of European scientific strength: if a field relies more on European science, citations to European journals would be higher.

B.2 Comparison Between Gap Measures

Table B1: PUBLICATIONS AND CITATIONS, EUROPE VS AMERICA

	Publications			Journal Citations		
OECD Subfield Equivalent	U.S.	Europe	Ratio	U.S. to U.S.	U.S. to Europe	Ratio
2.05 Materials Engineering	3,806	15,456	0.80	143	44	0.24
1.01 Mathematics	5,334	19,556	0.79	134	71	0.35
1.03 Physical Sciences and Astronomy	12,802	42,719	0.77	197	665	0.77
1.04 Chemical Sciences	31,330	75,596	0.71	650	656	0.50
3.02 Clinical Medicine	43,007	81,883	0.66	6,017	2,200	0.27
3.03 Health Sciences	5,121	9,373	0.65	1,042	336	0.24
4.01 Agriculture, Forestry, Fisheries	2,112	3,594	0.53	-	-	-
2.02 Electrical Eng, Electronic Eng	125	143	0.53	5	29	0.85
1.06 Biological Sciences	39,262	44,261	0.53	3,764	3,285	0.47
3.01 Basic Medical Research	32,556	34,614	0.52	4,845	2,721	0.36
2.01 Civil Engineering	1,010	636	0.39	-	-	-
1.05 Earth and Related Env Sciences	7,996	1,189	0.13	369	128	0.26

Notes: This table presents the number of citation-weighted articles (from WoS) that have non-missing subject and affiliation fields. The "Ratio" column for the Publications sub-columns divides the number of European-affiliated papers (published globally) divided by American-affiliated papers. The rows are downward-sorted by this value. The "Ratio" column for the Journal Citations sub-columns divides the number of citations to American journals by American journals by citations to European journals.

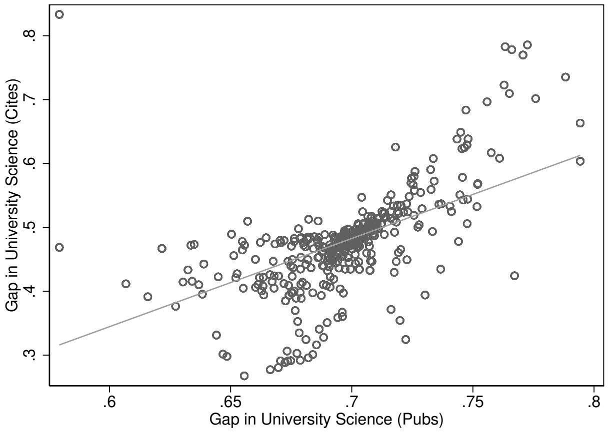
Table B1 compares the measures of scientific "strength" (relative backwardness). The "Ratio" columns for each measure present the number of European-authored papers and citations to European journals by American papers divided by the total number of papers and total number of citations by American journals, respectively. Intuitively, these ratios can be thought of as the "gap" or "lag" that exists between European and American institutions (fields with relatively large values are those where the scientific gap between Europe and the U.S. is large). The two measures do not yield identical results. Given the lack of citations data in civil engineering and agriculture, forestry & fisheries journals, the citations-based gap measure cannot be calculated for these fields. However, the fact that physics and chemistry have high gap scores, whereas clinical and medical sciences have relatively low gap scores, accords with the publications-based measure. A notable outlier in this measure is Electrical Engineering, which has a high score (0.85) partly due to low overall citations (34 citations in total throughout the 20-year period, compared to chemistry, which made 1,306 total citations).[59] Excluding this outlier, the corre-

---

[59] It is unclear whether this represents a measurement error. The only electrical engineering journal in print during this period (1900-20) is American ("Proceedings of the Institute of Radio Engineers") and the only other electrical engineering journal indexed in the SCI before 1940 is the BELL SYSTEM TECHNICAL JOURNAL, which is American. It is also possible that this field still relied on European science in the 1900-20 period, since 21 (62%) of the 34 citations were

lation between the citations-based measure and the publications-based measure is positive (r=0.286) at the scientific field level. At the firm level, i.e., when the observations are weighted by the industries and scientific subfields of firms in the sample, the correlation between the two measures (r=0.537) is greater, suggesting that fields with the highest mismatches between AMS and WoS are not very important in the patent classes used by our sample firms (Figure B3).

Figure B3: COMPARISON OF GAPS IN UNIVERSITY SCIENCE



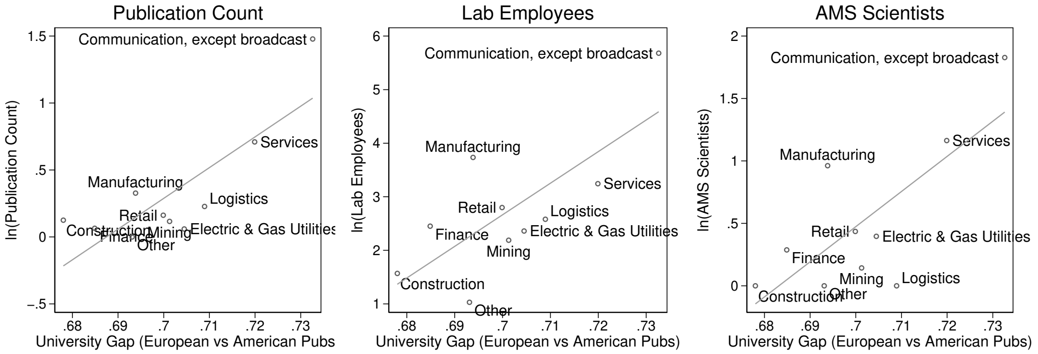
Notes: This figure compares the two scientific gap measures at the firm level. Higher values represent a larger gap between Europe and the United States. The journal citation-based gap measure (on the vertical axis) is positively correlated with the publication volume-based gap measure (on the horizontal axis) (r=0.537).

Appendix Figure B4 presents the correlation between the scientific gap and corporate science across industries. Corporate investments in science are greater in industries where the U.S. lags behind European science. For instance, construction, which relies on civil engineering, where the scientific gap is small, exhibits less corporate science investment than communications, which relies partly on chemistry, where the gap is large. This pattern is consistent with our conjecture in Section 2; it also calls for the use of industry fixed effects.

---

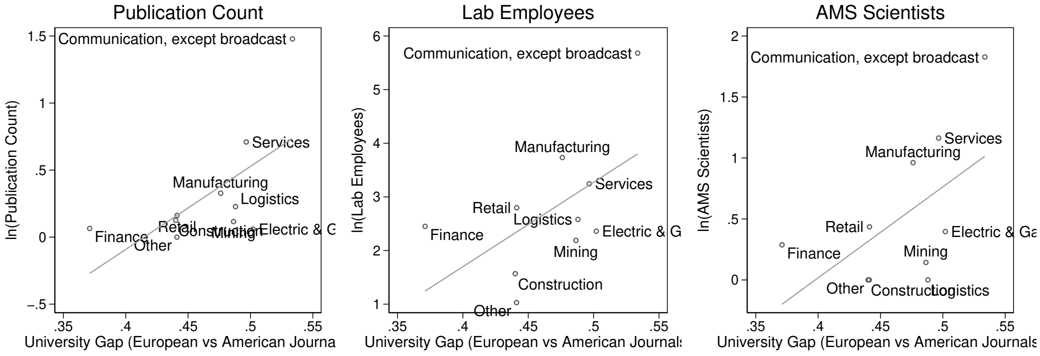
made to physics journals. Moreover, we have established in Section 2 that Electrical Engineering was a discipline where universities were unwilling or unable to provide scientific knowledge and published very little in. Hence, to the extent we measure American excellence in this discipline through the publication-based measure, it likely captures American *corporate* excellence, rather than university excellence.

Figure B4: CORPORATE SCIENCE VS GAPS IN UNIVERSITY SCIENCE, BY INDUSTRY



Notes: Industry-level scatter plots of firm investment in science and the gaps in the relevant academic discipline based on American and European publications in Web of Science. The left panel plots logged number of corporate publications per firm-year against gaps in university science (averaged at the 1-digit industry level). University science gap is measured as the ratio of European against American publications in Clarivate Web of Science. The middle panel replaces corporate publications with number of corporate lab employees; the right panel replaces it with scientists from AMS affiliated with firms.

Figure B5: CORPORATE SCIENCE VS GAP IN UNIVERSITY SCIENCE, BY INDUSTRY (CITATION-BASED GAP)



Notes: Industry-level scatter plots of firm investment in science and the gaps in the relevant academic discipline based on American citations to European journals in Web of Science. The left panel plots the natural log of one plus the publications per year against the gap measure. The middle panel replaces publications with the number of personnel at R&D labs, from the IRL directory, while the right panel replaces it with AMS scientists.

Appendix C Theory

C.1 Setup

There are three stages. In Stage 3, the firms compete in the product market. Their product market performance depends on the quality of their products and the cost of producing them. We assume that cost and quality depend upon the innovation output, d_i, $i = 0, 1$. Their payoffs from Stage 3 are $\Pi(d_0, d_1)$ and $\tilde{\Pi}(d_1, d_0)$, where the tilde indicates firm 1. We assume that $\Pi(d_0, d_1)$ is increasing in the first argument and decreasing in the second, and concave in its arguments, so that the firm's profit increases in its innovation output, albeit at a diminishing rate. To avoid the need for assumptions on third order derivatives, we assume

$$\Pi(d_0,d_1) = kd_0 - \frac{c_{00}}{2}d_0^2 - bd_1 - \frac{c_{11}}{2}d_1^2 + c_{01}d_1d_0,\ \ k > 1$$
$$\tilde{\Pi}(d_1,d_0) = d_1 - \frac{c_{00}}{2}d_1^2 - bd_0 - \frac{c_{11}}{2}d_0^2 + c_{01}d_1d_0$$

Firms farther from the frontier (e.g., smaller firms) can increase profits by imitation and by increasing scale, possibilities that the leaders have already exhausted. Instead, leaders have to introduce new and improved products and processes-to innovate. Accordingly, the marginal product of innovation for Firm 0 is greater than that of Firm 1 because $k > 1$.

The coefficient c_{01} is positive under strategic complementarity and negative under substitutability. Concavity of Π implies $c_{00} > 0, c_{11} > 0, c_{00}c_{11} - c_{01}^2 \geq 0$. We assume that $b > 0$ so that $\frac{\partial \Pi}{\partial d_1} = -b - c_{11}d_1 < 0$, i.e., innovation by rivals reduces payoff. We also assume that $c_{00} \geq c_{11}$. This assumption implies that the returns to internal invention increases at a slower rate than the rate at which profits decline due to invention by rivals.

In Stage 2, firms choose their innovation output. Firm 0 chooses d_0 and Firm 1 chooses d_1. The cost of innovation for Firm 0 is $\phi(r_0; u)d_0$, where r_0 represents investments in internal scientific research by the firm, and u indexes the stock of (relevant) public science. The cost of innovation includes the cost of inventing new products and processes or improving them. Internal research may directly lead to such inventions, but may also indirectly reduce the cost of invention by guiding the search for inventions in more promising directions. Innovations may also be based on inventions acquired from independent inventors, other firms or university researchers. Thus the cost of innovation also depends on the state of public science. It is natural to assume that both internal research and public science reduce the unit cost of innovation, $\phi(r_0;u)$, i.e., $\frac{\partial \phi}{\partial r_0} < 0, \frac{\partial \phi}{\partial u} < 0$, and diminishing returns so that $\frac{\partial^2 \phi}{\partial r_0^2} > 0$.

As we show below, the relationship between public science and internal research in the reduction in the unit cost of innovation will be important in how research investments relate to the stock of public science. The relationship may be one of complementarity (in the sense of Milgrom and Roberts 1989). For instance, it is typically believed that public science would complement internal research efforts. However, public science may also lead to startups and independent inventors, who can license or sell their inventions, which can substitute for internally generated inventions. If so, the relationship may be one of substitutability. Complementarity between university and corporate science exists if $-\frac{\partial^2 \phi}{\partial r_0 \partial u} > 0$, and substitutability exists if $-\frac{\partial^2 \phi}{\partial r_0 \partial u} < 0$. If $\frac{\partial^2 \phi}{\partial r_0 \partial u} = 0$, public science and research have independent effects on the cost of innovation.

The cost of innovation for Firm 1 is $\phi\tilde{(}u)d_1$. As noted, innovations may be based on external discoveries and inventions. Thus, we assume that $\phi\tilde{(}u)$ decreases with u.

In Stage 1, Firm 0 choose its research investments, r_0, and the cost of research is modelled simply as $\frac{\gamma}{2}r_0^2$, so $v_0 = kd_0 - \frac{c_{00}}{2}d_0^2 - bd_1 - \frac{c_{11}}{2}d_1^2 + c_{01}d_1d_0 - \phi(r_0,\lambda)d_0 - \frac{\gamma}{2}r_0^2$.

C.2 Stage 2: Innovation

We assume a stable Nash Equilibrium exists. For a stable equilibrium, we require that $D = c_{00}^2 - c_{01}^2 > 0 \iff |c_{00}| > |c_{01}|$.

Note that as long as $k \geq 1 + (\phi - \tilde{\phi})$, $d_0 \geq d_1$. In particular, if neither firm invests in research, so that $\phi = \tilde{\phi}$, Firm 0 would innovate more, and the gap is larger, the larger is k. This would imply that Firm 0 has a greater incentive to invest in research. The following intermediate results are helpful for later results.

C.2.1 Focal Firm Research and Innovation

The response of innovation output to the focal firm's research is

$$\begin{aligned} \frac{\partial d_0}{\partial r_0} &= \frac{c_{00}}{D}\left(-\frac{\partial \phi}{\partial r_0}\right) \\ \frac{\partial d_1}{\partial r_0} &= \frac{c_{01}}{D}\left(-\frac{\partial \phi}{\partial r_0}\right) \end{aligned} \tag{C2}$$

Note that if $c_{01} \geq 0$, Firm 1 also increases its innovation in response to an increase in research by Firm 0. Furthermore, $\frac{\partial^2 d_0}{\partial r_0 \partial u} = -\frac{c_{00}}{D}\frac{\partial^2 \phi}{\partial r_0 \partial u} \geq 0$ if $\frac{\partial^2 \phi}{\partial r_0 \partial u} \leq 0$, i.e., if public science and internal research are complements.

C.2.2 Public Science and Innovation

The response of innovation output to public science is

$$\begin{aligned} \frac{\partial d_0}{\partial u} &= \frac{-1}{D}\left(c_{00}\frac{\partial \phi}{\partial u} + c_{01}\frac{\partial \tilde{\phi}}{\partial u}\right) \\ \frac{\partial d_1}{\partial u} &= \frac{-1}{D}\left(c_{00}\frac{\partial \tilde{\phi}}{\partial u} + c_{01}\frac{\partial \phi}{\partial u}\right) \end{aligned} \tag{C3}$$

If there is strategic complementarity, i.e., $c_{01} \geq 0$, both firms innovate more in response to an increase in public science. However, if there is strategic substitutability, then one (but not both) firm may reduce innovation. In particular, if the innovation costs of a firm are not very responsive to public science, the effect of a rival increasing its innovation may cause the firm to reduce its innovation. However, note that

$$\frac{\partial d_0}{\partial u} + \frac{\partial d_1}{\partial u} = \frac{-1}{D}\left(\frac{\partial \phi}{\partial u} + \frac{\partial \tilde{\phi}}{\partial u}\right)(c_{00} + c_{01}) \geq 0 \tag{C4}$$

This implies that innovation on average increases with public science.

C.3 Stage 1: Research

Suppose Firm 1 does not invest in research. Firm 0 chooses r_0, taking into account how its choice will affect the equilibrium choices of d_0 and d_1 in the Stage 2 game. For Firm 0, the first-order condition for optimal r_0, is

$$-\frac{\partial \phi}{\partial r_0}d_0 + \frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial r_0} = \gamma r_0 \tag{C5}$$

The marginal return to research has a direct benefit represented by the first term: the reduction in the unit cost of innovation, which is proportional to the scale of innovation. The second term represents the feedback effect from competition in the innovation stage. By increasing innovation, research has a secondary benefit if it reduces innovation by the rival, which would be the case if there is strategic substitution in the innovation, so that $c_{01} \leq 0$. If innovations are strategic complements, then there is a secondary cost, because the second term would be negative. However, the first term is always larger than the second term. Substituting for $\frac{\partial d_1}{\partial r_0}$ from Equation C2 and gathering terms, Equation C5 can be rewritten as

$$-\frac{\partial \phi}{\partial r_0}\left(\frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D} + d_0\right) = \gamma r_0 \tag{C6}$$

Therefore, $\frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}+d_0$ must be positive at an interior maximum. A sufficient condition for this is strategic substitutability in innovation, $c_{01} \leq 0$. [60]

C.4 Innovation Leadership

Leaders earn higher profits. Conversely, the profits of the follower fall with the lead of Firm 0. Formally,

$$\begin{aligned} \frac{\partial v}{\partial k} &= d_0 + \frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial k} \\ &= d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D} > 0 \text{ at an interior maximum} \qquad \text{(C7)} \\ \frac{\partial \tilde{v}}{\partial k} &= \frac{\partial \tilde{\Pi}}{\partial d_0}\frac{\partial d_0}{\partial k} = \frac{\partial \tilde{\Pi}}{\partial d_0}\frac{c_{00}}{D} < 0 \end{aligned}$$

Importantly, the returns to research of the innovation leader increase with its lead k. Those of the follower decrease if innovations are strategic substitutes and increase otherwise. Intuitively, as k increases, the leader increases innovation. With strategic substitutes, the marginal return to innovation for the follower decreases. Given that research reduces the cost of innovation, the marginal return to research for the follower decreases.

$$\begin{aligned} \frac{\partial^2 v}{\partial k \partial r_0} &= \frac{\partial d_0}{\partial r_0} + \frac{c_{01}}{D}\left(-c_{11}\frac{\partial d_1}{\partial r_0} + c_{00}\frac{\partial d_0}{\partial r_0}\right) \\ &= \left(-\frac{\partial \phi}{\partial r_0}\right)\left(\frac{c_{00}}{D} + \frac{c_{01}^2}{D}(c_{00}-c_{11})\right) > 0 \\ \frac{\partial^2 \tilde{v}}{\partial k \partial r_1} &= \frac{c_{00}}{D}\left(-c_{11}\frac{\partial d_0}{\partial r_0} + c_{01}\frac{\partial d_1}{\partial r_1}\right) = \frac{c_{00}}{D}\left(-\frac{\partial \tilde{\phi}}{\partial r_1}\right)c_{01}(c_{00}-c_{11}) \leq 0 \iff c_{01} \leq 0 \end{aligned} \qquad \text{(C8)}$$

This result points to why the follower may not invest in research. Equation C8 implies that if innovations are strategic substitutes, as the gap between leaders and followers grows, their incentives to invest in research diverge: leaders are more likely to invest in research, and followers are less likely to do so. If there is a fixed cost to such investment, then, for a range of such costs, we will have only Firm 0 invest in research while Firm 1 does not.

C.5 Public Science

In this section, we focus on the equilibrium where only Firm 0 invests in research.

---

[60] We assume that the second order condition for an interior maximum holds. This requires that γ be large.

C.5.1 The Value of the Firm

The value of the firm, v, may decrease with public science if public science substitutes for internal research, particularly if innovations are strategic complements. Intuitively, although public science reduces the cost of innovation, the innovation cost of the rival also declines. Increased innovation by the rival reduces value for the focal firm. If public science substitutes for internal research, it will be less effective in reducing the innovation cost of Firm 0, i.e., $|\frac{\partial \phi}{\partial u}| < |\frac{\partial \tilde{\phi}}{\partial u}|$. Formally, the value of the firm is $v = \max_{r_0}\{\Pi - \gamma\frac{r_0^2}{2}\}$. Applying the envelope theorem, the effect of public science is given by

$$\begin{aligned}\frac{\partial v}{\partial u} &= -d_0\frac{\partial \phi}{\partial u} + \frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial u} \\ &= -\frac{\partial \phi}{\partial u}\left(d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}\right) - c_{00}\frac{\partial \Pi}{\partial d_1}\frac{\partial \tilde{\phi}}{\partial u}\end{aligned} \tag{C9}$$

Although the first term is positive by Equation C5, its magnitude depends on $|\frac{\partial \phi}{\partial u}|$. The second term is negative, and represents the effect due to the reduction in the rival's innovation cost. It is larger in magnitude the larger is $|\frac{\partial \tilde{\phi}}{\partial u}|$. Note that rivalry also matters. If $\dfrac{\partial \Pi}{\partial d_1} = -b + c_{01}d_0$ is large in magnitude (as would be the case for b large and $c_{01} < 0$), the firm's value can decline with public science.

C.6 Internal Research and Public Science

At an interior maximum, the direction of the effect of public science on internal research is given by $\dfrac{\partial^2 v}{\partial r_0 \partial u}$. Research increases with public science if $\dfrac{\partial^2 v}{\partial r_0 \partial u} \geq 0$ and decreases otherwise.

$$\begin{aligned}\frac{\partial^2 v}{\partial r_0 \partial u} &= \left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} + d_0\left(-\frac{\partial^2 \phi}{\partial r_0 \partial u}\right) + \frac{\partial \Pi}{\partial d_1}\frac{\partial^2 d_1}{\partial r_0 \partial u} + \frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u} \\ &\quad \text{substituting and collecting terms} \\ &= \left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} - \frac{\partial^2 \phi}{\partial r_0 \partial u}\left(d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}\right) + \frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u}\end{aligned} \tag{C10}$$

The first term in Equation C10 is positive. The second is positive if public science and research are complements in reducing the unit cost of innovation and negative otherwise. The third term is negative only if innovations are strategic complements and positive otherwise. Put differently, the first term reflects a direct effect: public science reduces innovation costs, and the resulting increase in innovation increases the marginal return to research. The second term represents the interaction between public science and research in reducing innovation costs. If they are complements, the second term also implies that the marginal return to research increases with public science. The third term captures the strategic interaction in innovation. If innovations are strategic substitutes, this term is also positive. Strategic complementarity is a necessary, but not sufficient, condition for this term to be negative. Thus, if internal research falls with public science, it implies that public science is a strategic substitute for research, or innovations are strategic complements, or both. These are one-way implications; even if they hold, public science could increase internal research if the direct effect, represented by the first term, is large.

To see this more fully, consider the case where there is neither complementarity nor substitution in the innovation stage, and where public science and research are independent in their effect on the unit cost of innovation. The latter implies that $\frac{\partial^2 \phi}{\partial r_0 \partial u} = 0$, and the former implies that $\frac{\partial d_1}{\partial r_0} = 0$. In that case, Equation C10 has a single term $\left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} \geq 0$. *That is, if public science and research are independent and there are no strategic interactions in the innovation stage, internal research increases with public science because public science increases the scale of innovation, thereby increasing the marginal return to research.*

If there are no strategic interactions in innovation, C10 is $\left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} - \frac{\partial^2 \phi}{\partial r_0 \partial u}$ $\left(d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}\right)$. The second term is non-negative if $-\frac{\partial^2 \phi}{\partial r_0 \partial u} \geq 0$, i.e., if public science and internal research are complements and negative otherwise. *Therefore, if internal research declines with public science, and there are no strategic interactions in innovation, it implies that public science and internal research are strategic substitutes.*

The third term can be written as

$$\frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u} = \frac{\partial d_1}{\partial r_0}\left[\frac{\partial^2 \Pi}{\partial d_1 \partial d_0}\frac{\partial d_0}{\partial u} + \frac{\partial^2 \Pi}{\partial d_1^2}\frac{\partial d_1}{\partial u}\right]$$

$$= \frac{\partial d_1}{\partial r_0}\frac{1}{D}\left[-c_{11}c_{00}(-\frac{\partial \tilde{\phi}}{\partial u}) - c_{11}c_{01}(-\frac{\partial \phi}{\partial u}) + c_{01}c_{00}(-\frac{\partial \phi}{\partial u}) + c_{01}^2(-\frac{\partial \tilde{\phi}}{\partial u})\right]$$

collecting terms and substituting

$$\frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u} = \frac{c_{01}}{D^2}(-\frac{\partial \tilde{\phi}}{\partial u})(c_{01}^2 - c_{00}c_{11}) + \frac{c_{01}^2}{D^2}(-\frac{\partial \phi}{\partial u})(c_{00} - c_{11}) \tag{C11}$$

Note that $c_{00} \geq c_{11}$, so that $\frac{c_{01}^2}{D^2}(-\frac{\partial \phi}{\partial u})(c_{00} - c_{11}) \geq 0$. Also, $-\frac{\partial \tilde{\phi}}{\partial u}(c_{01}^2 - c_{00}c_{11}) \leq 0$

by the concavity of Π. Thus, $\frac{c_{01}}{D^2}(-\frac{\partial\tilde{\phi}}{\partial u})(c_{01}^2 - c_{00}c_{11}) > 0$ if $c_{01} < 0$ and negative otherwise. Therefore, a necessary condition for the expression in C11 to be negative is that innovations be strategic complements. The conclusion is that for public science to reduce research, it would require that either innovations be strategic complements, or that public science be a strategic substitute for internal research. Else, public science will increase research by the leader.

C.6.1 The Gap Between the Leader and Follower, the Returns to Research, and Public Science

Recall from Equation C8 that the marginal returns from research to the leader as k increases is given by $-\frac{\partial\phi}{\partial r_0}\left(\frac{c_{00}}{D} + \frac{c_{01}^2}{D}(c_{00} - c_{11})\right)$. It is easy to see that this expression is increasing in u if public science and internal research are complements $(-\frac{\partial^2\phi}{\partial r_0 \partial u} \geq 0)$ and decreasing otherwise.

That is, restricting ourselves to the case where only the leader invests in research, we have that

$$\frac{\partial r_0}{\partial k} = -\frac{\partial\phi}{\partial r_0}\left(\frac{c_{00}}{D} + \frac{c_{01}^2}{D}(c_{00} - c_{11})\right)\left(-\frac{\partial^2 v}{\partial r_0^2}\right)^{-1} > 0 \tag{C12}$$

The effect of public science u on $\left(-\frac{\partial^2 v}{\partial r_0^2}\right)^{-1}$ cannot be signed in general. However, the term $-\frac{\partial\phi}{\partial r_0}\left(\frac{c_{00}}{D} + \frac{c_{01}^2}{D}(c_{00} - c_{11})\right)$ will increase with public science if $-\frac{\partial^2\phi}{\partial r_0 \partial u} \geq 0$ and decreasing otherwise. This suggests that if public science and internal research are substitutes, firms closer to the technological frontier will respond to decreases in public science by increasing internal research. Put differently, suppose $\left(-\frac{\partial^2 v}{\partial r_0^2}\right)^{-1}$ is constant. Then, $\frac{\partial^2 r_0}{\partial k \partial u} \geq 0 \iff -\frac{\partial^2\phi}{\partial r_0 \partial u} \geq 0$.

Table C2: EFFECT ON INTERNAL RESEARCH

Conceptual Relationship	Analytical Relationship	Empirical Measure	Univ Science Complementary $(-\frac{\partial^2\phi}{\partial r_0 \partial u} > 0)$	Univ Science Substitute $(-\frac{\partial^2\phi}{\partial r_0 \partial u} < 0)$
Proximity to Frontier	$\frac{\partial r_0}{\partial k}$	$k \left\{ \begin{array}{r} \text{Patent Cites to Science} \\ \text{First Patent in CPC} \\ \text{Patent Importance} \\ \text{Market Share} \end{array} \right.$	+	+
Univ Science	$\frac{\partial r_0}{\partial u}$	$u \left\{ \begin{array}{c} \text{Eur. Scientific Pubs} \\ \text{Cites to Eur. Journals} \end{array} \right.$	+ if Strategic Substitutes	Ambiguous
Proximity to Frontier and Univ Science	$\frac{\partial^2 r_0}{\partial k \partial u}$	Same As Above	+	-

Appendix D Auxiliary Results

Table D1: BREAKTHROUGH INVENTIONS AND RELIANCE ON SCIENCE (OLS)

Dependent Variable	Top 10% KPST		Top 5% Cites	
	(1) 1900-1919	(2) 1920-1940	(3) 1900-1919	(4) 1920-1940
Dummy for patent citation to science	0.000	0.122	0.092	0.054
	(0.083)	(0.021)	(0.113)	(0.014)
Avg of Dep Var	0.053	0.105	0.051	0.055
Year Dummies	Yes	Yes	Yes	Yes
4-digit CPC Dummies	Yes	Yes	Yes	Yes
R^2	0.090	0.379	0.029	0.023
Observations	672,342	848,358	672,944	848,826

Notes: The unit of analysis is the patent. Sample is limited from 1900 to 1940 and split into two periods (1900-1919 for Columns 1 and 3 and 1920-1940 for Columns 2 and 4). The dependent variable in Columns 1 and 2 is a dummy indicating whether the patent is within the top 10% of the patent text-based importance measure from Kelly, Papanikolaou, Seru, and Taddy (2021). The dependent variable for Columns 3 and 4 is a dummy indicating whether the patent's forward patent citations are within top 5% of the sample. Standard errors are robust to arbitrary heteroscedasticity.

Appendix Table D1 shows that, even after controlling for year and patent class, patents citing science are more likely to be breakthrough in the 1920-1940 period (Columns 3 and 4), whereas no such relationship exists for the 1900-1919 period (Columns 1 and 2).

D.1 Replications of Level Results

Table D2: Robustness Tests for Investment in Science and Technological Frontier

	Probit			Firm Ownership			Top Assets Quintile Subsample			Exclude Outliers	Between-Firm	
Dependent Variable	(1) Pub Dummy	(2) Lab Dummy	(3) AMS Dummy	(4) Pub Count	(5) Lab Empl	(6) AMS Scientists	(7) Pub Count	(8) Lab Empl	(9) AMS Scientists	(10) Pub Count	(11) Pub Count	(12) Lab Empl
Dummy for Patent Cite to Science	0.425	0.835	-0.042	1.259	0.571	1.203	1.264	0.649	1.405	1.552	0.287	0.453
	(0.179)	(0.357)	(0.410)	(0.528)	(0.357)	(0.437)	(0.697)	(0.375)	(0.541)	(0.167)	(0.352)	(0.279)
Dummy for First Patent in CPC	0.347	0.126	1.909	0.715	0.595	1.733	0.847	0.550	2.101	0.936	2.399	2.368
	(0.102)	(0.142)	(0.620)	(0.258)	(0.337)	(0.358)	(0.278)	(0.224)	(0.497)	(0.134)	(0.539)	(0.491)
Patent Importance	1.069	0.562	1.571	4.576	0.181	0.897	2.784	0.098	5.283	2.365	4.317	0.380
	(0.343)	(0.393)	(0.718)	(1.534)	(2.006)	(1.188)	(1.360)	(0.470)	(1.436)	(0.471)	(1.296)	(0.951)
ln(Assets)	0.297	0.206	0.771	0.762	0.606	0.275	0.981	0.614	-0.049	0.315	0.519	0.413
	(0.057)	(0.058)	(0.137)	(0.094)	(0.129)	(0.168)	(0.195)	(0.183)	(0.175)	(0.060)	(0.118)	(0.110)
Distance to Universities	0.000	0.000	-0.000	0.000	-0.000	0.000	0.001	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Dummy for Business Group				-0.646	-0.129	0.746						
				(0.174)	(0.351)	(0.301)						
Ultimate Owner Sales HHI				-1.127	-0.273	0.362						
				(0.464)	(0.953)	(0.506)						
Average of Dep Var	0.161	0.612	0.391	2.260	295.817	4.892	3.414	252.347	9.551	0.446	0.543	50.168
Year Dummies	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.270	0.185	0.400	0.735	0.705	0.709	0.731	0.686	0.730	0.536	0.738	0.804
Number of Firms	397	338	202	155	93	111	85	71	49	388	389	401
Number of Obs	2,640	907	202	388	93	111	587	222	49	2,536	389	401

Notes: The analysis is at the firm-year level for Columns 1-2, 4-5, 7-8, 10, and at the firm level for Columns 3, 6, 9, 11-12. “ln(Assets)” takes the natural log of concurrent assets when unit of analysis is at the firm-year; of average assets between 1926-1930 where dependent variable is AMS Scientists (Columns 3,6,9); and of average assets over the full sample period for Columns 11 and 12. Columns 1-3 estimate a Probit regression where the dependent variables are equal to one if a firm publishes (Column 1), operates a lab (Column 2), or employs AMS scientists (Column 3). Columns 4-12 estimate a conditional Poisson regression with publication, lab employee and AMS scientist counts as dependent variables. The sample in Columns 4-6 is limited to observations in the B&M sample for which there are sales data. Further details on constructing Business Group and Ultimate Owner Sales HHI are in Appendix A.3. The sample in Columns 7-9 is limited to firms in the top quintile of assets. The sample in Column 10 excludes firms above the 95 percentile of total publications and above the 95 percentile of average asset size over the sample period. Standard errors are clustered at the firm level. Please see Appendix Table A1 for details on variable construction.

In addition to extending the analysis in Table 3, Table D2 adds several robustness checks: Since size may also be correlated with attributes of the firm's ownership structure, we control for whether the firm is part of a business group (with at least three public affiliates) in Columns 4-6 of Table D2 and find that the positive correlation between investment in science and firm size continues to hold (Appendix A.3 provides details on the construction of firm ownership data).

Following Nelson (1959), we also test whether diversified firms are more likely to invest in science in Columns 4-6 and find mixed results: the HHI coefficient estimates (measuring the diversification of sales across industries calculated for each ultimate owner's controlled firms) for Columns 4 and 5 are consistent with diversification being positively correlated with corporate science, but the signs are reversed in Column 6.

Table D3: CORPORATE SCIENCE AND LEADERSHIP: SATURATED SPECIFICATION (POISSON)

Dependent Variable	Publication Count	Lab Employees	AMS Scientists
	(1)	(2)	(3)
Dummy for Patent Cite to Science	2.572	1.042	1.701
	(0.438)	(0.342)	(0.436)
Dummy for First Patent in CPC	1.334	1.271	1.264
	(0.403)	(0.203)	(0.560)
Patent Importance	-0.423	0.923	1.132
	(1.160)	(0.529)	(0.514)
ln(Assets, 1926-1930)	0.331	0.511	0.248
	(0.195)	(0.076)	(0.133)
Market Share	2.787	-0.461	1.752
	(1.213)	(0.608)	(0.805)
Dummy for Competitive Market	-1.359	-0.801	-0.772
	(0.568)	(0.266)	(0.407)
Distance to Universities	0.001	-0.000	-0.000
	(0.000)	(0.000)	(0.001)
Average of Dep Var	1.137	105.332	2.822
Year Dummies	Yes	Yes	No
3-Digit SIC Dummies	No	No	No
Pseudo-R^2	0.667	0.647	0.660
Number of Firms	272	241	197
Number of Obs	2,371	798	197

Notes: The analysis is at the firm-year level for Columns 1-2, and at the firm level for Column 3. "ln(Assets, 1926-1930)" takes the natural log average assets between 1926-1930. Standard errors are clustered at the firm level. Please see Appendix Table A1 for details on variable construction.

Table D3 estimates a conditional Poisson specification which includes all level variables from Tables 3 and 4.

D.2 Replications of Results with Alternative Measure of University Gap

Tables D4, D5, and D6 replicate Tables 5, 6, and 7 by replacing the default (publication-based) gap measure with the journal citation-based gap measure.

Table D4: CORPORATE SCIENCE AND JOURNAL CITATION-BASED GAPS IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	Publication Count (1)	Lab Employees (2)	AMS Scientists (3)
University Gap (Cites)	3.550	0.964	2.853
	(2.881)	(1.595)	(2.147)
ln(Assets)	1.019	0.840	0.948
	(0.251)	(0.081)	(0.152)
Distance to Universities	-0.001	-0.000	-0.001
	(0.001)	(0.000)	(0.001)
Average of Dependent Variable	0.688	61.267	2.261
Year Dummies	Yes	Yes	No
3-Digit SIC Dummies	Yes	Yes	Yes
Pseudo-R^2	0.617	0.663	0.545
Number of Firms	422	394	276
Number of Observations	3,855	1,331	276

Notes: The analysis is at the firm-year level for Columns 1 and 2 and at the firm level for Column 3. “ln(Assets)” takes the natural log of concurrent assets for Columns 1 and 2, and of average assets between 1926 and 1930 for Column 3. “University Gap (Cites)” divides the number of American journal citations to European journals by all citations made by American journals between 1900 and 1920. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table D5: CORPORATE SCIENCE AND JOURNAL CITATION-BASED GAPS IN UNIVERSITY SCIENCE, BY TECHNOLOGICAL LEADERSHIP (POISSON)

Dependent Variable	Publication Count			Lab Employees			AMS Scientists		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
University Gap (Cites)	-1.323	-0.757	1.440	-0.946	-0.059	2.908	-0.824	-3.380	1.207
	(1.804)	(2.280)	(3.063)	(2.432)	(1.787)	(1.916)	(1.611)	(1.909)	(2.218)
× Dummy for Patent Cite to Science	32.132			14.048			21.529		
	(7.032)			(5.280)			(4.327)		
× Dummy for First Patent in CPC		13.115			3.793			17.379	
		(4.096)			(2.340)			(3.908)	
× Patent Importance			61.951			15.087			38.595
			(15.331)			(7.275)			(10.823)
Dummy for Patent Cite to Science	-13.680			-5.526			-8.177		
	(3.481)			(2.334)			(1.978)		
Dummy for First Patent in CPC		-4.678			-0.788			-5.621	
		(2.021)			(1.027)			(1.738)	
Patent Importance			-24.375			-5.492			-14.877
			(7.050)			(2.961)			(4.803)
ln(Assets)	0.747	0.690	0.938	0.696	0.663	0.761	0.621	0.555	0.880
	(0.106)	(0.146)	(0.122)	(0.067)	(0.071)	(0.076)	(0.124)	(0.130)	(0.160)
Distance to Universities	0.000	0.000	-0.001	-0.000	-0.000	-0.000	0.000	0.000	-0.001
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Average of Dependent Variable	0.688	0.688	0.939	61.267	61.267	82.962	2.261	2.261	2.886
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.724	0.683	0.713	0.713	0.700	0.685	0.683	0.698	0.620
Number of Firms	422	422	397	394	394	338	276	276	202
Number of Observations	3,855	3,855	2,640	1,331	1,331	934	276	276	202

Notes: The analysis is at the firm-year level for Columns 1-6 and at the firm level for Columns 7-9. “ln(Assets)” takes the natural log of concurrent assets for Columns 1-6, and of average assets between 1926-1930 for Columns 7-9. “University Gap (Cites)” divides the number of American journal citations to European journals by all citations made by American journals between 1900 and 1920. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

Table D6: CORPORATE SCIENCE AND JOURNAL CITATION-BASED GAPS IN UNIVERSITY SCIENCE, BY MARKET LEADERSHIP (POISSON)

Dependent Variable	Publication Count			Lab Employees			AMS Scientists		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
University Gap (Cites)	-25.009	-2.510	6.966	-32.874	-0.324	4.059	13.681	-1.327	3.720
	(48.482)	(2.110)	(1.613)	(29.986)	(2.198)	(1.974)	(43.432)	(2.249)	(2.154)
× ln(Assets, 1926-1930)	1.049			1.201			-0.394		
	(1.791)			(1.115)			(1.609)		
× Market Share		69.266			26.629			35.868	
		(29.061)			(18.481)			(15.972)	
× Dummy for Competitive Market			-8.161			-7.429			-8.187
			(4.072)			(3.070)			(5.381)
ln(Assets, 1926-1930)	0.662			0.330			1.143	0.759	0.868
	(0.903)			(0.575)			(0.849)	(0.199)	(0.096)
Market Share		-28.428			-12.745			-15.275	
		(14.753)			(8.876)			(7.621)	
Dummy for Competitive Market			2.926			2.821			2.916
			(1.853)			(1.417)			(2.409)
Distance to Universities	-0.001	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
ln(Assets)		0.634	0.919		0.863	0.880			
		(0.078)	(0.109)		(0.127)	(0.078)			
Average of Dependent Variable	0.628	0.922	0.624	56.157	78.168	58.817	2.261	2.789	1.768
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
3-Digit SIC Dummies	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Pseudo-R^2	0.623	0.740	0.342	0.627	0.688	0.529	0.545	0.579	0.370
Number of Firms	355	273	417	355	260	389	276	209	354
Number of Observations	4,860	2,747	4,252	1,720	934	1,386	276	209	354

Notes: The analysis is at the firm-year level for Columns 1-6 and at the firm level for Columns 7-9. "ln(Assets)" takes the natural log of concurrent assets; "ln(Assets, 1926-1930) takes average assets between 1926 and 1930. Market Share" averages firm market share at the 3-digit industry level between 1926 and 1930. "University Gap (Cites)" divides the number of American journal citations to European journals by all citations made by American journals between 1900 and 1920. Please see Appendix Table A1 for details on variable construction. Standard errors are clustered at the firm level.

D.3 Replications of Firm Performance Results

Table D7: CORPORATE SCIENCE AND STOCK MARKET VALUE (CITATION-BASED GAP) (OLS)

Dependent Variable	ln(Market Capitalization)			
	University Gap (Cites) Split		University Gap (Cites) Split for Subsample with Lab & AMS Scientist	
	(1) Univ Gap Below Mean	(2) Univ Gap Above Mean	(3) Univ Gap Below Mean	(4) Univ Gap Above Mean
ln(Publication Stock$_{t-1}$)	0.239	0.212	0.184	0.268
	(0.090)	(0.072)	(0.128)	(0.101)
ln(Patent Stock$_{t-1}$)	0.211	0.016	0.186	0.056
	(0.043)	(0.043)	(0.071)	(0.040)
ln(Assets$_{t-1}$)	0.724	0.794	0.990	0.698
	(0.074)	(0.068)	(0.134)	(0.062)
Average of Dependent Variable	19.567	19.248	20.457	20.472
Year Fixed Effects	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes
R^2	0.727	0.798	0.860	0.918
Number of Firms	167	154	45	32
Number of Observations	1,776	1,557	515	368

Notes: The analysis is at the firm-year level. The dependent variable is logged market capitalization. Columns 1 and 2 split the sample by mean values of the "University Gap (Cites)" measure based on share of American journal citations to European journals. Columns 3 and 4 limit the sample to firms that operate a lab and employ AMS scientists. Standard errors are clustered at the firm level.

Table D8: CORPORATE SCIENCE AND MARKET-TO-BOOK RATIOS (OLS)

Dependent Variable	ln(Tobin's Q)				
		University Gap Split		University Gap (Cites) Split	
	(1) All	(2) University Gap Below Mean	(3) Univ Gap Above Mean	(4) Univ Gap Below Mean	(5) Univ Gap Above Mean
Publication Stock/Assets$_{t-1}$	6.860	-3.483	10.926	4.610	5.610
	(4.247)	(4.199)	(4.076)	(8.110)	(4.690)
Patent Stock/Assets$_{t-1}$	0.150	0.157	0.175	0.107	0.406
	(0.088)	(0.165)	(0.103)	(0.081)	(0.139)
Average of Dependent Variable	-0.479	-0.517	-0.440	-0.462	-0.498
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.350	0.437	0.415	0.395	0.459
Number of Firms	316	165	151	164	152
Number of Observations	3,213	1,625	1,586	1,716	1,497

Notes: Unit of analysis is at the firm-year level. Dependent variable is log of Tobin's Q. Columns 2 and 3 split the sample by mean values of the "University Gap" measure comparing European and American publications. Columns 4 and 5 split the sample by mean values of the "University Gap (Cites)" measure based on citations to European journals. Year and industry dummies at the 2-digit SIC code level are included in all columns. Standard errors are clustered at the firm level.

References for Appendices

Berle, A., & Means, G. (1932). Private Property and the Modern Corporation. *New York: Mac-millan*.

Bonbright, J. C., & Means, G. C. (1932). *The holding company: Its public significance and its regulation*. McGraw-Hill Book Company.

Buchanan, N. S. (1936). The Public Utility Holding Company Problem. *Calif. L. Rev.*, *25*, 517.

Flagg, C. A. (1913). *A List of American Doctoral Dissertations Printed in 1912*. United States Government Printing Office.

Graham, J. R., Leary, M. T., & Roberts, M. R. (2015). A Century of Capital Structure: The Leveraging of Corporate America. *Journal of Financial Economics*, *118*(3), 658–683.

Iaria, A., Schwarz, C., & Waldinger, F. (2018). Frontier knowledge and scientific production: Evidence from the collapse of international science. *The Quarterly Journal of Economics*, *133*(2), 927–991.

Kandel, E., Kosenko, K., Morck, R., & Yafeh, Y. (2019). The Great Pyramids of America: A Revised History of US Business Groups, Corporate Ownership and Regulation, 1926-1950. *Strategic Management Journal*, *40*(5), 781–808.

Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, *3*(3), 303–20.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, *132*(2), 665–712.

Lundberg, F. (1937). *America's 60 Families*. Vanguard Press New York.

Marx, M., & Fuegi, A. (2020). Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles. *Strategic Management Journal*.

Moser, P., & Kim, S. (2022). Women in Science. Lessons from the Baby Boom.

Moser, P., & Parsa, S. (2022). McCarthy and the Red-ucators: Effects of Political Persecution on Science.

Moser, P., Parsa, S., & San, S. (2022). Immigration, Science, and Invention. Lessons from the Quota Acts. *Available at SSRN 3558718*.

Nelson, R. R. (1959). The Economics of Invention: A Survey of the Literature. *The Journal of Business*, *32*(2), 101–127.

Petralia, S., Balland, P.-A., & Rigby, D. L. (2016). Unveiling the Geography of Historical Patents in the United States from 1836 to 1975. *Scientific Data*, *3*(1), 1–14.

Pinchot, G. (1928). The Power Monopoly; Its Make-up and Its Menace.

Wilcox, C. (1940). Competition and Monopoly in American Industry. *Investigation of Concentration of Economic Power, T.N.E.C, Monograph No. 21*.

Wilson, M. M. (1932). *A List of American Doctoral Dissertations Printed in 1930*. United States Government Printing Office.