NBER WORKING PAPER SERIES

THE RISE OF SCIENTIFIC RESEARCH IN CORPORATE AMERICA

Ashish Arora
Sharon Belenzon
Konstantin Kosenko
Jungkyu Suh
Yishay Yafeh

Working Paper 29260
http://www.nber.org/papers/w29260

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2021, Revised November 2024

We thank seminar participants at Bar Ilan University, Ben Gurion University, Boston University, the Hebrew University, Northwestern 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, the 2023 GSU-CEAR Finance Conference, the 2023 NYU Economics of Strategy Conference, as well as Naomi Lamoreaux, Tim Simcoe, Bhaven Sampat, Tarun Khanna, and Filippo Mezzanotti 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, Tamara Calzado Real, Sophie Cote, and Hansen Zhang 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 views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research or the Bank of Israel.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w29260

The Rise of Scientific Research in Corporate America
Ashish Arora, Sharon Belenzon, Konstantin Kosenko, Jungkyu Suh, and Yishay Yafeh
NBER Working Paper No. 29260
September 2021, Revised November 2024
JEL No. N8, N82, O32

ABSTRACT

It is widely believed that university and corporate research are complementary: Companies invest in research in part to develop the capacity to absorb the knowledge emerging from universities. However, as we show in this paper, corporate research in the United States emerged when American universities were behind the world frontier in scientific research. Why, then, did for-profit businesses choose to invest in creating new knowledge, much of which could spill over to rivals, and whose conduct presented many managerial challenges? We argue that corporate research in America arose in the 1920s to compensate for weak university research, not to complement it. 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 were unable to meet their needs. Large firms, close to the technological frontier, and operating in concentrated industries were likely to invest in research, especially in scientific disciplines where American universities lagged behind the scientific frontier. Corporate science seems to have paid off, resulting in novel patents and high market valuations for firms engaged in research.

Ashish Arora
Fuqua School of Business
Duke University
Box 90120
Durham, NC 27708-0120
and NBER
ashish.arora@duke.edu

Sharon Belenzon
Fuqua School of Business
Duke University
100 Fuqua Drive
Durham, NC 27708
and NBER
sharon.belenzon@duke.edu

Konstantin Kosenko
Bank of Israel
Jerusalem 91907
Israel
konstantin.kosenko@boi.org.il

Jungkyu Suh
Stern School of Business
New York University
44 West Fourth Street
New York, NY 10011
js13220@stern.nyu.edu

Yishay Yafeh
School of Business Administration
The Hebrew University, Mt. Scopus
Jerusalem 91905
Israel
and CEPR and ECGI
msyafeh@huji.ac.il

1 Introduction

The systematic application of scientific knowledge is a key source of technological advancements and economic growth. A substantial literature has emerged to study university-to-industry knowledge transfer (Bercovitz & Feldman, 2008; Bikard & Marx, 2019; Cohen et al., 2020; Jaffe et al., 2007; Perkmann et al., 2021; Rothaermel et al., 2007; Zucker et al., 2002) and public-private partnerships to address technological challenges (Agarwal et al., 2021; Babina, He, et al., 2023; Gross & Sampat, 2020a; Howell, 2017). At the heart of this research is an assumption that companies invest in research to access and to absorb the knowledge emerging from universities (Cohen & Levinthal, 1989), or, as Nathan Rosenberg, referring to why firms invest in basic research, put it, "...basic research may be thought of as a ticket of admission to an information network" (Rosenberg, 1990, p. 170).

This paper documents that corporate research in the United States began when American universities were behind the world frontier in scientific research. Though American firms did innovate before the 1920s, they performed very little R&D, and no internal scientific research, sourcing new inventions from independent inventors. This changed gradually as firms not only invested considerable amounts in R&D, but some also invested in scientific research (Mowery, 2009). And while some American universities had begun to engage in research, their investments were dwarfed by their corporate counterparts.

Why, then, did for-profit businesses choose to invest in creating new knowledge, much of which could spill over to rivals, and whose conduct presented many managerial challenges? This paper argues that corporate research in American firms arose in the 1920s to compensate for inadequate university research, particularly of the type relevant for the needs of firms on the technological frontier. The argument is as follows: Inventions were increasingly based in science. However, independent inventors, hitherto the primary source of inventions to firms, lacked the requisite scientific expertise (Lamoreaux et al., 1999; Nicholas, 2010). American universities, though a valuable source of trained engineers and scientists, were unwilling or unable to carry out the use-inspired scientific research required to address the technological needs of industry. The lack of commercially relevant research induced large firms, such as GE, AT&T, DuPont, and Kodak to invest in internal research. Holding large market shares and operating in concentrated industries, such companies did not fear spillovers to competitors.

We study these issues empirically using a newly assembled, comprehensive historical dataset of firms operating in 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. For each firm in the sample, we assemble three proxies for investment in research — scientific journal publications (a research output), the operation of research labs, and the employment of prominent scientists (research inputs). 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. 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. To measure the public science that firms could draw upon, we develop new measures of relative American scientific backwardness, in comparison with Europe, by scientific field. Our sample period, ranging from 1926 to 1940, is well suited to study our research question because it is when corporate science emerged in the United States. Moreover, in contrast to the period following the Second World War, the U.S. government during our sample period exerted little influence on corporate research and development activities (Gross & Sampat, 2020b; Mowery & Simcoe, 2002).

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 provide evidence that American universities did not produce scientific knowledge, relevant and at the levels needed by the leading firms. Consistent with (Bikard, 2018), we show that university-authored papers during our period were less likely to be cited by patents or to be textually close to them compared to papers authored by corporate scientists. A case study of early electronics research provides qualitative support for the proposition that American universities failed to support the scientific advances made before the Second World War. We find that firms investing in science relied on scientific fields in which U.S. universities were relatively weak at the time (such as select sub-fields of chemistry and physics), suggesting that corporate research aimed to compensate for, rather than complement, university science.

Third, we establish that firms at the technological frontier invested in scientific research: Firms with patents citing scientific articles (our proxy for proximity to the technological frontier) published nearly 38 times more in scientific journals and employed fifteen-fold more lab personnel than firms whose patents did not cite science.

Fourth, we provide evidence that 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. Thus, they could afford to carry out research, a classic public good, without fearing spillovers to rivals.

Finally, we observe a positive correlation between the intensity of corporate research, the generation of valuable patents, and high stock market valuations, consistent with investment in corporate research being a source of competitive advantage.

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

We contribute to the existing literature in several ways. First, we add a novel explanation for why firms invest in science. Our argument, that corporate research in America was a response to a need for science coupled with deficiencies in public science, departs from previous theories that focus on absorptive capacity (Cohen & Levinthal, 1989; Rosenberg, 1990), recruitment incentives (Sauermann & Cohen, 2010), and product quality signaling (Azoulay, 2002; Hicks, 1995). While we are not the first to argue knowledge substitution exists between academia and industry, we are the first, to our knowledge, to show that this substitution will be carried out principally by large and technologically advanced firms, rather than smaller startups. Though histories of prominent companies are consistent with our argument (Hounshell, 1996; Hounshell & Smith, 1988; Jenkins & Chandler, 1975; Maclaurin & Harman, 1949; Reich, 1985), to our knowledge, no work has systematically investigated it using economy-wide data on corporate innovation.

Our work is also related to the question of why many U.S. firms have substantially reduced their investments in scientific research in recent decades (Arora et al., 2018; Bhaskarabhatla & Hegde, 2014; Mowery, 2009). Our findings suggest that increased funding of university research in general (particularly in fields such as solid-state electronics and computer science) starting after World War II may have eventually reduced the private returns to scientific research, possibly leading to a decline in corporate science in subsequent decades (Arora et al., 2018).

Our results also speak to the institutional voids framework in a context hitherto not documented in the literature. Khanna (2000), and Khanna and Palepu (2000) show that diversified business groups fill such voids in emerging markets. Our findings indicate that when public institutions, such as university research, are weak, then building internal research capabilities can be a source of competitive advantage, especially for firms with the scale and scope to justify the fixed cost.

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, et al., 2023) and scientists (Moser et al., 2022). This new resource should open new research on the potential links between corporate research and development, and government policy.

Finally, our use of historical methods to complement the quantitative analysis joins recent works in organization studies (Braguinsky & Hounshell, 2016; Bucheli & Wadhwani, 2014; Ingram et al., 2012; Silverman & Ingram, 2017) that employ hermeneutics (probing the motivations of the decision maker), contextualization (interpreting the transpired events in their respective time and space), and source criticism (weighing the integrity of the surviving evidence) to extend their explanatory power beyond their settings (Pillai et al., 2024). We draw on biographical sources on managers and corporate scientists to corroborate our thesis that universities were failing to provide science relevant to firms. We also contextualize our findings in the backdrop of rapid scientific advances in physics and chemistry, which offered opportunities to technologically advanced firms for breakthrough innovations. For the present-day context, our findings speak most directly to sectors experiencing significant technological opportunities created by advances in related scientific fields. Recognizing that our qualitative evidence relies on the more successful firms in the period, we assemble an economy-wide dataset of innovating firms to examine our thesis more systematically.

The paper is organized as follows. Section 2 surveys the historical context; Section 3 describes our data; Section 4 presents our econometric specifications and estimation results; Section 5 concludes. Appendices present more details on the historical data, a theoretical framework and empirical robustness tests.

2 The Rise of Corporate Science in America

American firms began investing in scientific research around the 1920s owing to three factors. First, inventions increasingly relied on science. Second, some firms reached the technological frontier, having exhausted the opportunities to innovate through trial-and-error. Third, industrially relevant public science from universities was inadequate. The following historical survey illustrates these points.

A Growing Corporate Need for Science — The beginning of the twentieth century was marked by a burst of scientific advances in chemistry and physics 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 First World War rarely cited the scientific literature. The following period, between 1920 and 1940, on the other hand, shows a marked increase in citations to science, especially by breakthrough patents.[1] Breakthrough chemical patents, 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.

[Figures 1 and 2 Here]

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 133 scientists listed in the *American Men of Science* directory of prominent scientists, compared to only 18 in 1900.[2] 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 substantial growth in scientific publications by corporations from 1900-1919 to 1920-1941: Scientific publications grew by more than 11 fold for chemical firms and over 29 fold for firms in electrical engineering between these two periods.

Abundant scientific opportunities, coupled with a scarcity of publicly available scientific knowledge relevant for industrial application, encouraged 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 to invest in science internally. Company histories indicate that, in the early years, corporate labs focused on quality control and solving operational problems rather than fundamental science.[3] As innovation became more science-based, companies initially looked to external suppliers and specialized contract research organizations, such as the Mellon Institute, established in 1913.[4] But contract research worked best for generic, well-specified, problems, where contracting problems were less severe, and required that the firm itself possess significant

[1]Breakthrough patents are patents in the top decile of the "importance" measure from Kelly et al. (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.

[2]The American Men of Science (AMS) directory is a comprehensive listing of prominent scientists in the United States since 1906 (Moser et al., 2022). Further details on the dataset are discussed in Section 3.

[3]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). Their research in surface chemistry (for Langmuir) and crystal physics and electronic diffraction (for Davisson) were not only scientifically important but was also relevant to their companies.

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

scientific capabilities (Mowery, 1983). Contract research could complement internal research, not substitute it.

[Table 1 Here]

Public Science was Inadequate for Industry — Universities were another potential source. However, American universities were unwilling or unable to produce the required science. Column 1 of Table 1 regresses citations received from patents against whether the paper is authored by a corporation. The estimates from the linear probability model suggest that university papers are around 90% less likely to be cited by a patent than a corporate paper. Even after adding dummies for scientific subfield and journal, university papers are 76% less likely to be cited than corporate papers (Column 2). In Columns 3 and 4, we test whether university publications are farther from patents than corporate publications by comparing textual similarity.[5] Though universities produced the vast bulk of publications, the coefficient estimate in Column 3 show that university publications are 34% less likely than corporate publications to be ranked among the 10 closest papers to a patent.[6] Adding subfield and journal fixed effects in Column 4 reduces this gap to 6.4%, but the difference is still statistically significant. That subfield and journal dummies reduce coefficient estimates indicates that universities also publish in different subjects and outlets from companies. Figure 3 compares the focus of corporate and university publications. We find that electrical engineering, physics, and chemistry account for the top three corporate scientific subjects, with electrical engineering accounting for over half of all corporate publications between 1926 and 1940. In contrast, the top three university fields are clinical medicine (28%), biology (23%), and chemistry (9%).

[Figure 3 Here]

The lower proportion of university science that is commercially relevant (i.e., science that is readily applied to inventions) notwithstanding, the absolute quantity may still be high; 92% of all science & engineering papers published between 1926 and 1940 were by university authors. We therefore examine how often university science is relied upon compared to corporate science. We find that fewer than half (38%) of science-citing patents cite public (non-corporate) science (5% of patent citations are to papers by nonprofits,

[5] Textual similarity is calculated using the SPECTER algorithm, which uses a transformer-based neural network to process texts. Details on its construction are provided in (Arora et al., 2023).

[6] There are in total 39,616 papers: 3,363 are by companies and 36,253 are by universities. Of the corporate papers, 3,215 (95.6%) have at least one patent for which they are ranked as the top 10 most similar papers. For university papers, 22,302 (61.5%) are ranked within the top 10 similarity score of any patent.

7% to those by the government, 26% to those by universities).[7] Moreover, "breakthrough" patents (defined as the top 5% most important patents based on patent textual similarity measures from Kelly et al. (2021)) have around 9% lower university science citation rates (24%) compared to non-breakthrough ones (33%).[8]

The evidence also supports the argument that university science received less attention from technologically consequential inventors than corporate science. This is consistent with modern evidence from Bikard (2018) who finds that, even within paper "twins" on simultaneous discoveries, academic papers are less likely than corporate papers to be cited by inventors in the life sciences. This is not necessarily because American university science during our sample period was inferior to firms: university papers received around 28% more 5-year forward citations from other peer-reviewed scientific papers (1.40 vs 1.09). Rather, as Figure 3 showed, U.S. university research before World War II was concentrated in scientific pursuits less relevant to industrial needs, leading American corporations on the technology frontier to invest in research themselves. The case 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 discovery of the electron in 1897 by J.J. Thomson built upon earlier research in electricity and magnetism by Maxwell and Hertz. 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 emission," the phenomenon would form the basis of the radio industry and vacuum tube electronics in the coming decades.

The immediate impact, however, was more prosaic – extending the life of the incandescent bulb. General Electric (GE) had exhausted trial and error methods to reduce the blackening that occurred on the surface of the incandescent light bulb. The competing hypothesis to Richardson's theory argued that thermionic

[7]There are 190 patents making citations to 123 papers, which constitute 203 citation pairs.

[8]To augment citation-based measures of patent-paper proximity, we examine the share of university papers textually close to a patent in Appendix Table D2. Columns 1 shows that more important patents are closer to corporate publications relative to university publications. Using a stricter measure of similarity strengthens this association sixfold, as Column 3 shows. Column 2 shows that higher quality patents (forward citations) are closer to corporate publications relative to university publications, and more so when using a stricter standard for closeness (Column 4).

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 impurities (such as water vapor) left inside the lamp, which led to prescriptions to create a better vacuum inside the bulb (Birr, 1957; Langmuir, 1913; Soddy, 1907). Irving Langmuir's first research project at GE Research Labs was to settle this debate, a classic example of commercially relevant science. Langmuir, an American chemist hired by GE after completing his PhD at the University of Göttingen in Germany, attempted a solution suggested by Richardson's theory: instead of trying to create a better vacuum, Langmuir proposed to fill the bulb with inert gases that would scatter the evaporated particles.[9]

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, notably the diode (1904 by John Ambrose Fleming) and triode (1906 by Lee De Forest). 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.

2.1.2 Only the market-leading firms seemed willing to invest in electronics research

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. 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 a larger output. Beyond scale, these firms also urgently needed improvements in their core product: 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 also created new opportunities in vacuum tubes (Birr, 1957).[10]

Unlike GE, Westinghouse neither had the scale nor the urgent need to improve the incandescent bulb. Westinghouse only had 13% of the lamp market. In addition, the antitrust settlement of 1911 whereby GE

[9]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).

[10]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 (licensed to Westinghouse for $ 1 million in 1894), and Leo Arons' mercury vapor lamp.

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).[11]

A similar story emerges in telephony. As telephony became more science-based, AT&T realized it could not rely upon independent inventors and needed internal scientific capability. In 1900, AT&T 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). But 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. De Forest himself 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 Francisco in 1915.[12]

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. Western Union had pioneered innovations such as Ezra Cornell's glass-insulated telegraph wires in the 1840s but stayed out of the telephone market as part of an 1879 patent settlement with Bell. It also largely ignored the opportunities from advances in electrical engineering and chemistry.[13]

2.2 Undersupply of Industrially Relevant Public Science in Electronics

Despite the evident need, American universities were unable or unwilling to invest in research in electronics at the scale required. Maclaurin and Harman (1949) argue that much of the scientific foundation in electronics until the 1920s was provided by European institutions such as the Cavendish laboratory in Britain

[11] 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).

[12] 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.

[13] AT&T employed 26 scientists listed in the AMS directory of 1921, while Western Union employed none.

(Faraday, Maxwell, Thomson, Richardson, and Bragg) and German research universities (Hertz, Siemens, Hittorf, Roentgen, Wehnelt, and Braun).[14] Until the First World War, American universities were well behind: They 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).

While university scientists were instrumental in establishing the American Physical Society (1889 at Columbia University), and the American Chemical Society (1876 at New York University), societies studying commercially relevant topics were founded by corporate scientists: the American Institute of Electrical Engineers, established in 1884, was headed by Norvin Green from Western Union. The first president of the Institute of Radio Engineers, established in 1912, was Robert Marriott from the Wireless Company of America. Similarly, the Optical Society was founded in 1916 by Peter Nutting of Kodak, while the Acoustical Society of America was founded by Harvey Fletcher and V.O. Knudsen in 1928 (Bell Telephone Laboratories).

The absence of American universities from electronics research can be attributed to a combination of a i) general reluctance (by the government) to fund research, ii) the high cost of electronics research as well as iii) a distaste among faculty for industrially-relevant, applied subjects. 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 thousand per university on research (National Research Council, 1938). The University of Chicago ($2.6 million in 1929-30), and the University of California ($2.4 million 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 million, respectively (Maclaurin & Harman, 1949, p.158).[15] Finally, the intellectual climate at American universities discouraged academics from engaging with industry: Henry Rowland's (Johns Hopkins) 1883 in his address

[14]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.

[15]Research equipment was expensive. For instance, The generation of high vacuum — a pre-requisite to studying thermionic emissions — required Langmuir to adopt and modify Wolfgang Gaede's molecular pump imported from Germany. The ultra-centrifuges, Svedberg's Nobel prize-winning scientific instrument used to separate chemical substances by their molecular weights was vital to polymer research. In the 1920s, apart from DuPont, only the University of Wisconsin (where Svedberg had visited in 1923) had an ultra-centrifuge (Cerveaux, 2013).

to the American Association for the Advancement of Science (AAAS), titled “A Plea for Pure Science” stated that he was “tired of seeing our professors degrading their chairs by the pursuit of applied science instead of pure science” (Kline, 1995; Rowland, 1883).

2.3 Generalizing from the Case Studies

The case study of electrical engineering suggests that firms began investing in scientific research for three reasons: i) inventions increasingly relied upon science, and sometimes directly arose from scientific breakthroughs; ii) some firms reached the technological frontier, having exhausted the opportunities to innovate through trial-and-error; and iii) industrially relevant public science was inadequate.

We formalize this intuition in an analytical model in Appendix A involving a leader firm and a laggard firm. We distinguish between upstream research and downstream innovation. Research reduces the cost of innovation to the firm, as does public science. Thus, the returns to investing in research depend on the scale of innovation - the larger the scale, the greater the marginal return to innovation. Leaders, with a larger scale of innovation, are therefore more likely to invest in research. Weaker public science could create a competitive advantage for market leaders who invest in science. Weaker public science increases the cost of innovation. Unlike the market leader, however, the rival is unable to offset this through internal research, thereby increasing the gap between it and the market leader. In our empirical analysis, we investigate whether investment in scientific research is associated with higher market value and whether this is conditioned by public science.

The response to an increase in public science depends on whether public science increases the efficiency of internal research, and whether the firm’s innovations are strategic complements or substitutes for those of its rival (which can also access public science). The marginal returns to the leader of investing in research are greater when public science is weak if public science substitutes for internal research. A necessary condition for public science to reduce internal research is either that public science substitutes for internal research, or that the marginal return to innovation is lower when a rival increases innovation (strategic substitutes). In our empirical analysis, we examine whether market leadership and proximity to the technical frontier are associated with greater investment in internal research and whether this relationship is conditioned by public science.

3 Data

Our unbalanced firm panel 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 major "industrial" (i.e., non-financial) firms found in Berle and Means (1932) (hereafter B&M) and their subsidiaries.[16] We only include firms that patent at least once between 1926 and 1940 in an IPC that cites at least five scientific articles between 1947 and 1957, to focus on firms that are "at risk" of engaging in scientific research.[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 augment the B&M sample with 235 additional listed firms from CRSP that patent in science-citing IPCs (Kogan et al., 2017). Our basic sample thus consists of 466 American firms, both private and public, that patent at least once between 1926 and 1940 in an IPC that cites at least five scientific articles between 1947 and 1957. For these, there are 4,282 firm-years for which we have financial statement data between 1926 and 1940.

3.1 Corporate Investment in Science

Scientific Publications — We source 283,992 peer-reviewed scientific papers (excluding humanities and the social sciences) from Microsoft Academic Graph (MAG) published between 1926 and 1940. We match 140,766 author affiliations from these publications to our sample firm names using fuzzy string-matching, accounting for abbreviations common in the era such as RR for railroad, and name variants such as SOCONY for the Standard Oil Company of New York.[18] We exclude articles authored by eponymous charitable foundations and hospitals (e.g., by DuPont, Carnegie and Rockefeller). 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.[19]

While publications were not the primary purpose of firms, they were tolerated to entice the brightest

[16] 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.

[17] 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.

[18] 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.

[19] See Appendix B.1 for details on matching scientific publications to firms.

researchers to join their ranks. University scientists were often reluctant to join firms, leading to some being offered the freedom to publish, in addition to a salary premium. A biography of Willis Whitney, GE's first research lab manager, also suggests that the "(F)reedom to do fundamental research and to publish the results had attracted scientists of the caliber of (William) Coolidge and (Colin) Fink" (Wise, 1985)[20] C.E.K. Mees, the founding head of Kodak's research laboratory, negotiated the right to publish. He also encouraged company researchers not to neglect the exploration of attractive problems of purely theoretical interest. He believed that even applied scientists needed "the constant stimulation of theory ", which could additionally yield unexpected but valuable industrial applications (Clarke, 1971). Wallace Carothers was offered a substantial pay raise (from $3,200 at Harvard vs $5,000 at DuPont) and the freedom to work on scientific areas of interest to him, with substantial research support in terms of new assistants and a new laboratory (Hounshell & Smith, 1988).

Reich (1985, p.251) highlights a strategic reason as well (Shvadron, 2024): publications would encourage outside researchers to work on the same areas as well:

> [Through publications], the areas of research interesting to industry came forcefully to the attention of the larger scientific and technical communities, causing researchers outside the industrial labs to consider some of the issues raised within them. The development of concern with the electron physics of vacuum-tube operations is an excellent case in point. By the early 1920s there was, as one physicist noted, a "fever of commercialized science" sweeping the American scientific community.

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 (Service, 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.

American Men of Science (AMS) — We use the employment of prominent scientists by a firm as an additional indicator of corporate investment in science. For this, we collect information on the employment affiliation of American scientists from the 1921 and 1938 edition of the Cattel Directory of American Men of Science.

[20] [First names and capitalization added by authors.

As seen in Appendix Figure B2, corporate investment in science has a skewed distribution: a total of 308 firms (66%) 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.

[Table 2 Here]

The three indicators of corporate investment in science are highly correlated.[21] However, each measure is imperfect. Table 2 shows that nearly half of all firms that operate research labs do not publish in the sciences, and around 37% do not employ 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.

3.2 University Science

3.2.1 Gap in University Science

Measuring the gap between the required and available industrially relevant science is very difficult. Instead, we use as a proxy the "relative backwardness" of American academia compared to Europe by scientific field. Intuitively, if the same fraction of university science is industrially relevant in both Europe and the United States, and if firms could not easily access scientific knowledge from Europe, then our proxy would capture differences across sectors in the extent to which relevant science was available to industry. Indeed, knowledge flows from Europe were restricted after the onset of World War I.[22] Our primary measure is based on the citation-weighted number of scientific publications authored by scientists in each region. We find similar results using an alternative measure based on citations to European journals made by American scientific journals.

Scientific Publications: U.S. and Europe — We use the country of correspondence for the authors of scientific publications. We collect the author address 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. For publications with missing addresses, we match the authors' last and first names to the AMS directory to identify 27,924 publications by prominent American scientists. The

[21] 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.

[22] Iaria et al. (2018) show that World War disrupted scientific knowledge transfer from Europe, which failed to recover as late as 1930.

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.[23] The above process yields 155,571 scientific publications by Europeans and 60,605 publications by Americans between 1900 and 1920. We weigh the publication counts by the number of forward paper citations received until 2019.

Constructing Firms-Specific Scientific Gaps — Our regional scientific activity data from Web of Science are encoded at the scientific field level. Therefore, we link 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.[24] 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.[25] 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.[26]

3.2.2 University Human Capital

Geographical proximity to universities could affect the access a firm has to university science and human capital (Furman & MacGarvie, 2007). For each firm we calculate the average distance between its headquarters' location (from the 1929-1930 edition of Moody's Manuals) and American universities granting graduate degrees in the natural sciences in 1930. For firms with no information on headquarter location, we match their patents to the HistPat dataset, which extracts Federal Information Processing System (FIPS) County codes for patent assignees (Petralia et al., 2016). We impute the firm's address as the one that

[23] We use the correspondence in Marx and Fuegi (2020) to map Web of Science subject fields to 39 OECD subfields. Appendix Table C1 provides the breakdown by field.

[24] 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).

[25] The weight is the share of patents filed by the firm in the IPC, between 1926 and 1940.

[26] 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 C1, 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 C presents a detailed example, showing each step of the calculation.

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.[27]

We also measure the number of PhD graduates that are relevant to a firm's technological base by counting the number of dissertations similar to a firm's patent portfolio for a focal year. We calculate textual similarities (ranging from zero to one) between dissertations (titles and abstracts) sourced from ProQuest Dissertation and Theses (PQDT) Database to the focal firm's patents.[28] We sum these similarity scores at the firm-year level to derive the weighted number of PhD graduates relevant to a firm.

3.3 Patents

Patent data are derived from Google's public patent dataset. There are 637,190 patents granted between 1926 and 1940. We match 92,330 patents to the 466 firms in our panel between 1926 and 1940 (see Appendix Section B.1 for details). We normalize forward patent citations by dividing the total number of citations received by a focal patent by the per-patent citations received by all patents granted in that year.

3.3.1 Proximity to the Technological Frontier

We identify firms close to the technological frontier in three ways: first, we measure the extent to which its patents "rely" on science by counting citations to scientific publications in Microsoft Academic Graph in the text of the patent, from Marx and Fuegi (2020).[29] Second, we measure the "novelty" of a patent by measuring whether it is the first in its Cooperative Classification Class (CPC). Third, 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.

3.4 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

[27] Appendix Section B.2 provides further details.

[28] When a dissertation is similar to multiple patents by a firm, we keep the highest similarity score.

[29] We use in-text citations because NPL citations are available only after 1947. We use references with a confidence score of 9 or above. We find 237 patent citations to science by our sample firms between 1926 and 1940.

Moody's Manuals, and fill in data for the intervening years from the same source.[30] 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 by dividing firm sales in a year by 3-digit industry sales in the same year.[31] We then average annual market share for the first five years of the panel (1926-1930) for each firm.[32] We also use Wilcox (1940) to classify 3-digit industries as competitive and non-competitive.[33]

3.5 Descriptive Statistics

The maximal number of observations is 6,990 (466 firms observed over 15 years between 1926 and 1940).[34] There are only around a third of the total observations for "Lab Employees" because the IRL data were collected for only five sample years by the NRC (1927, 31, 33, 37, 40).

[Figure 4 Here]

Figure 4 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 to i) publish, ii) have a lab, or iii) 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, 27.6% of the 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 other patent-based measures of proximity to

[30] We use additional, ownership-related variables, drawn from Moody's Manuals, in a robustness test in Table D6

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

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

[33] Wilcox uses a broad set of criteria to separate competitive and non-competitive industries for the period 1934-1939. Stigler (1949) and Nutter and Einhorn (1969) both validate the Wilcox (1940) classification.

[34] Appendix Table B2 presents descriptive statistics at the firm-year level, and Appendix Table B1 provides variable definitions.

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 measure are less pronounced).

[Table 3 Here]

Table 3 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 4 specifications with each of these measures separately, and all three of them together.

4 Empirical Results

4.1 Which Firms Invest in Science?

Table 4 presents the results of conditional Poisson specifications regressing corporate investment in science on measures of proximity to the technological frontier and market leadership, controlling for firm size (assets), geographical distance to universities, and the number of PhD dissertations relevant to firm patents.[35] All specifications control for year and industry fixed effects and standard errors are clustered at the firm level.

[Table 4 Here]

Corporate publications, our primary indicator of investment in science, are positively associated with proximity to the frontier – patent citations to science, first-in-CPC dummy, and patent importance – individually (Columns 1-3) and jointly (Column 4). The estimate from Column 1 implies that citing scientific articles in patents is associated with 8.5 times more corporate publications. Appendix Tables D3 and D4 replicate these specifications for lab size and the number of prominent scientists: Patent citations to science are correlated with 2.5 times more lab employees and 5.5 times more AMS scientists. In Appendix Table D5, we find that the result in Table 4 Column 1 is principally driven by firms citing recent and highly cited science. The estimate in Column 1 implies that firms whose patents only cite papers older than average are not more likely to publish papers themselves.[36] Firms with patent citations to recent science publish 20% more than those with citations to any scientific article. Column 2 shows that firms whose patents cite "high quality"

[35]Geographical distance to universities and number of PhDs dissertations proxy for the human capital the firm has access to.

[36]A paper is "Recent" if the difference between its publication year and the citing patent's grant year is below the average for the CPC-grant year.

science (papers with above average forward citations from other scientific articles) publish twice as much as those that cite any scientific article.[37]

Columns 5-8 of Table 4 focus on the relationship of corporate science with absolute size (assets), relative size (market share) and industry concentration (competition).[38] All three are correlated significantly with corporate investment in science, with a standard deviation higher market share associated with 2.6 times more publications (Column 6).

We also test whether proximity to the technological frontier can predict the decision to invest in science at all. Appendix Table D6, Columns 1-3 show that firms with scientific patents are around 49% more likely to publish (Column 1), 38% more likely to operate labs (Column 2) and 68% more likely to employ a prominent scientist (Column 3). 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 Appendix Table D6 and find that the positive correlation between investment in science and firm size continues to hold (Appendix B.3 provides details on the construction of firm ownership data).[39] We also find that our results tend to hold for the largest firms in the sample (Columns 7-9), and when the largest firms are excluded (Column 10).

4.1.1 American Universities And Industrially Relevant Science

We explore the extent to which the relative weakness of American academia in certain fields (i.e., the "gap" in public science) accentuates the incentives of the affected firms to invest in research. Column 1 of 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, number of relevant American PhD dissertations, as well as year and industry dummies. The point estimate implies that a standard deviation higher gap in university science is associated with around twice as many corporate publications.[40] In Column 2, we interact distance to universities with gap measures and find that the baseline results are stronger for firms located close to universities: the marginal effect of *University Gap* on corporate publications in Column 1 becomes negative for firms above the 95th percentile of distance to universities (and statistically indistinguishable

[37] The results are similar for lab size (Columns 3 & 4) and employment of prominent scientists (Columns 5 & 6).

[38] Since these measures are correlated with technological frontier measures (Table 3), we include them altogether in Appendix Table D7 and find the main results hold. Note that we measure assets and market share 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.

[39] 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 Column 4 is consistent with diversification being positively correlated with corporate science, but estimates in Columns 5 and 6 are not statistically significant.

[40] The results are qualitatively similar for lab employees and AMS scientists in Column 1 of Appendix Tables D8 and D9.

from zero for firms above the 90th percentile). Firms farthest from universities would have been least likely to benefit from academic knowledge, which could explain why the gap matters less for them.

[Table 5 Here]

Columns 3-5 interact the university gaps with proximity to the technological frontier.[41] The interaction coefficient estimate in Column 3 suggests that the positive association between the scientific gap and investment in corporate research is driven by firms whose patents cite science. Comparing university gap measures at the 25th and 75th 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). Columns 4 and 5 are also consistent with the view that the scientific gap matters more for firms closer to the technological frontier.

Columns 6-8 interact university gaps with firm size, market share and concentration. The interaction term estimate in Column 7 implies that the correlation between gap and publications is around 2.8 times larger for firms in the 75th percentile of market share compared to those in the 25th percentile. Column 8 similarly shows that the positive association between scientific publications and the university gap is principally driven by firms in non-competitive industries.[42] We also see similar results using lab employees and AMS scientists (Columns 6-8 of Appendix Tables D8 and D9). We replicate these results using an alternative measure of university gap that is based on the number of citations from American journals to European ones for publications (Appendix Table D11), lab personnel (Table D12), and prominent scientists (Table D13).[43] In sum, the firms most likely to respond to gaps in science were technologically advanced and faced limited competition.

4.2 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 of such inventions: 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 (in the top 5% of their respective

[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.

[42] We find the interaction term with respect to size (assets) in Column 6 to be statistically insignificant, which may indicate that market share and competition capture appropriability from scientific knowledge that is distinct from size.

[43] Details on constructing this gap measure are found in Appendix Section C

[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.

measures for each grant year) against investment in science and controls for lagged size (assets) and patent stock. Note that patent value is conceptually distinct from patent-based measures of proximity to the technological frontier used in Tables 4, 5, and D6. Patents that cite science, are first in their technology class, or are "important" may actually be discounted by investors and future inventors *because* of their novelty.

[Table 6 Here]

The estimates from Table 6, Columns 1 & 2 show that corporate publications are positively correlated with the number of highly valuable patents. An increase in publication stock by one standard deviation is associated with about 16% more patents in the top 5% of stock market value. We replicate these results for lab size as well the employment of prominent scientists in Appendix Table D10. 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 Columns 3-9 (Belenzon, 2012). A standard deviation larger publication stock (normalized by assets) is associated with a 0.7% increase in logged firm market capitalization relative to the sample mean (Column 3). However, splitting the sample by university gap measures in Columns 4 and 5, we find that firms whose gaps in university science are above the sample mean (Column 5) are driving the association between market value and publication stock.[45]

Columns 6 and 7 show that publication stock is positively associated with stock market value primarily for firms that employ AMS scientists and have a lab (Column 7). Firms in Column 6, which do not engage in both activities, show a 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 8 and 9 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 43% stronger for firms whose fields are related to scientific disciplines in which American universities lagged behind Europe.[46] 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

[45] Appendix Table D14 shows a similar result using citations to European journals as an alternative measure of university gap. Appendix Table D15 replicates the specification for Tobin's Q.

[46] While the 95% confidence intervals for Columns 6 and 7 as well as 8 and 9 overlap, it is also clear that the Column 6 and 8 estimates are small and imprecise (i.e., not significantly different from zero) while the Column 7 and 9 estimates are larger and more precise.

science.

4.3 Alternative Explanations

The results so far are consistent with American corporations investing in scientific research to make up for insufficient university knowledge during this period. We consider several alternative explanations that may explain our empirical patterns.

Absorbing European Science — It is possible that firms were investing in internal research to absorb scientific knowledge from Europe. This may explain why firms in fields with greater gaps to Europe tend to invest in science, especially if they are at the technological frontier (Table 5). Indeed, leading firms employed scientists trained in Europe, and backward publication citations to European journals are positively correlated with the employment of these scientists.[47] However, university science from Europe, while advanced, was also removed from application, leaving gaps in translational research to be filled by firms. For instance, U.S. patent citations to Europe are rare, with only 19 out of 203 papers cited (9.3%) by corporate patents having been published in European journals.

German research universities and public research institutions were at the forefront of the chemical sciences, with German chemists accounting for 44% of Nobel Laureates before 1920. However, this did not prevent Bayer, BASF, Hoechst or AGFA from establishing their own basic research programs. German academics were initially sources of inventions as well as human capital during the late nineteenth century (Murmann & Landau, 1998). Carl Graebe and Carl Lieberman at the University of Berlin, for instance, collaborated with Heinrich Caro of BASF to produce Alizarin in 1869. Hoechst's first drug, antipyrin, was a product of a Joint Venture with Ludwig Knorr at University of Erlangen around the 1880s. Carl Duisberg (who went on to direct Bayer Research and later became CEO) started his research career in 1883 at Strasbourg University by carrying out basic research on areas of interest to Bayer. By the end of the 1880s, however, German firms determined that they could no longer rely upon universities. They funded industrial research consortia such as the Kaiser Wilhelm Institutes for Chemistry, Physical Chemistry and Coal Research (Murmann & Landau, 1998, p.38-39). However, many of the leading firms also established central laboratories: BASF (1879), Hoechst (1883), Oehler (1882), Kalle (1882), Agfa (1882), and Bayer (1886) (Homburg, 1992).

[Table 7 Here]

[47] Appendix Figure D2 plots citations to European journals by firm publications against the employment of European PhDs and finds a positive correlation (r= 0.245)

To support our claim that American firms were investing in science not to absorb European science, but to generate their own research, we compare the research topics of doctorate holders in AMS (1921) by the sector they work in (corporate or university) and their PhD institution (U.S. versus European). To measure research similarity, we measure the Jaccard similarity between the words used in the topic section.[48] There are 15,988 unique topics and 4.4 topics per scientist. Of the 5,496 scientists, 356 have a corporate affiliation at least once during their career, and 22% (1,200) had European doctoral training. To mitigate concerns that we are capturing industry's (and academia's) propensity to publish in different disciplines (e.g., industry scientists may focus on chemical engineering, while academia focuses on chemistry), we limit our comparison to scientists within the same field (there are 121 fields).

Table 7 shows that the average topic similarity between corporate scientists and university scientists with European PhDs (2.2%) is 12% lower than corporate similarity with American PhDs (2.5%). If corporate investments in science were primarily aimed at absorbing European science, we would expect corporate scientists' topics to be closer to those with European PhDs, contrary to our findings.

[Figure 5 Here]

Other evidence suggests that European scientists changed research towards commercially relevant directions after moving to American firms. We analyze the publications of 747 corporate scientists in AMS (1938) before and after joining a firm for the first time in their career. Figure 5 shows that scientists' publications receive 2.4 times more citations from patents 10 years after joining a firm (compared to the 10 years prior), and the pattern holds for European-trained scientists.[49] Corporate scientists also author around 30% more publications, while their average publication citations grow almost nine-fold between the two periods. The evidence points toward corporations investing in science (rather than merely invention). It also points to corporate science being qualitatively different (focused in practical applications), yet respected by the scientific community. In sum, while European science may have been more advanced compared to the U.S. in key areas, it was still insufficient for the needs of American corporations.

Reverse Causality — A second alternative explanation is that universities may be responding to an abundance of corporate research rather than firms responding to inadequate university research. The estab-

[48]The Jaccard index is calculated by dividing the count of topics shared between two scientists by the total count of distinct topics across both scientists, represented as $\frac{A \cap B}{A \cup B}$ for the research topics of scientists A and B. Further details can be found in Appendix Section B.1.4.

[49]There are 50 (6.7%) such scientists, which shows these labs' scientific ranks were primarily staffed by American-trained personnel. Further details on matching AMS scientists to scientific publications are in Appendix Section B.1.4.

lishment of Caltech, for instance, was partly a reaction to the perceived excessive industry partnerships advocated by the MIT Technology Plan in the 1920s (Geiger, 1986). That is, it is possible that universities vacated fields where firms had established dominance early. Though we cannot rule this out, it seems unlikely that universities stepped back from research in areas where corporations had invested. We calculate our gap measures prior to the sample period, between 1900 and 1920. As established in Figure 2, this is a period when corporate investment in science was very low (with corporate publications in chemistry and electrical engineering averaging only around three papers per year). Moreover, As noted earlier, funding for university research was limited to begin with. For instance, Langmuir joined GE in 1909, in part, because his academic job (at the Stevens Institute of Technology) offered him little time and resources for research. Similarly, Clint Davisson left Carnegie Tech and joined Western Electric in 1917 because it offered him more research freedom. It seems more likely that the leading companies stepped into scientific topics that universities were neglecting but that the companies valued.

5 Discussion and Conclusion

The twentieth century saw technical progress becoming 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 materials and instruments they were using. For large firms that were close to the technological frontier and operated in concentrated industries, internal research was a way to overcome the inability or unwillingness of American universities to provide the needed knowledge and thereby gain a competitive advantage: competitors would not be able to acquire the needed scientific knowledge from universities.

As he was making the case for a central research laboratory for "pure science or fundamental research work" to the executive committee of DuPont in 1926, Charles Stine argued that "applied research is facing a shortage of its principal raw materials" (Hounshell & Smith, 1988, p.223). The study of large molecules would eventually open a new field in chemistry specializing in polymers and spawn the plastics and synthetic fiber industry. However, American universities during this time were not investing in this topic. This led DuPont to recruit Wallace Carothers from Harvard university, who would publish his landmark paper on polymerization in 1931, and the firm was selling nylon stockings in the New York World Fair in 1939, extending its market position in the chemical industry.

A similar process seems to be at work in select scientific subfields such as the birth of semiconductor physics in the 1950s and the renaissance of Artificial Intelligence (AI) in the 2010s. In the former, AT&T invested heavily in the creation of a solid-state alternative to the vacuum tube, which led to the transistor in 1948. Similar concepts had been floated by academics such as Julius Lilienfeld in the 1920s, but no follow-on work at universities ensued, prompting the firm to invest (Riordan & Hoddeson, 1997). Similarly, university support for AI research into neural networks using deep propagation had dried up in all but a few institutions during the 1990s (Goodfellow et al., 2016). However, after seeing the promise of deep learning from computer vision competitions such as ImageNet (principally, Alexnet in 2012), Google invested heavily in AI research, hiring university researchers such as Geoffrey Hinton (University of Toronto, founder of DNN Research) and integrating them into their research organization to work on internal projects such as the Google Neural Machine Translation tool in 2016. Both cases demonstrate that technologically leading firms in concentrated markets can expand the frontiers of knowledge when university investments are anemic.

On the other hand, our results also suggest that these leading firms, while substituting for the lack of academic science, also reinforced their advantage over competitors. This is especially relevant to the recent debates over rising industry concentration in the United States, and on the merits of regulation of "big tech" firms in the ICT sector(Autor et al., 2020; Gutiérrez & Philippon, 2020). Our results imply that the breakup of firms with substantial research capabilities may, at least in the short term, reduce industrial research, as the demise of the Bell Labs showed. Policymakers should therefore also weigh the unique functions that "big tech" firms provide — the provision of commercially relevant scientific research neglected by the public sector — against the downsides of their market power.

Our findings have implications for answering whether and why our innovation system has become more sclerotic (Bloom et al., 2017). Following World War II, the growth of university research was paralleled by growing investments in corporate research. However, by the 1980s, the two trends had begun to diverge, with universities continuing to spend more on research and corporations withdrawing from it (Arora et al., 2020). Our results suggest that sustained federal support for universities, as well as reforms that have reduced the barriers to transferring government-funded science at universities (such as the Bayh-Dole Act of 1980), may have made it easier for firms to buy or borrow commercially relevant science rather than produce it themselves. How this growing division of labor between academia and universities is related to the rate and direction of technological change is a question we leave for future research.

We acknowledge that the relationship between academic and industrial research is complex: universities not only produce knowledge that feeds into firms' innovations but also train scientists and engineers. Universities also produce inventions that compete with corporate inventions. While we control for heterogeneity in firm access to human capital by proximity to universities and the availability of relevant doctoral dissertations in the regression analyses, further research is needed to unpack how public science, in all its dimensions, is related to private innovation.

In addition to providing new evidence on corporate research in America in the interwar period, we assemble the most extensive historical sample of American firms 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. Though, to our knowledge, there is no corresponding dataset for European firms, a natural follow-on research agenda may examine how European firms responded to the availability of public science during the same period. We hope that these newly developed data will spur future research on the open questions we have raised.

References

Agarwal, R., Kim, S., & Moeen, M. (2021). Leveraging Private Enterprise: Incubation of New Industries to Address the Public Sector's Mission-Oriented Grand Challenges. *Strategy Science*, *6*(4), 385–411.

Arora, A., Belenzon, S., Cioaca, L., Sheer, L., & Zhang, H. (2023). The Effect of Public Science on Corporate R&D. *Working Paper*.

Arora, A., Belenzon, S., & Patacconi, A. (2018). The decline of science in corporate R&D. *Strategic Management Journal*, *39*(1), 3–32.

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.

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. (2023). Financial Disruptions and the Organization of Innovation: Evidence from the Great Depression. *The Review of Financial Studies*, 2023.

Babina, T., He, A. X., Howell, S. T., Perlman, E. R., & Staudt, J. (2023). Cutting the Innovation Engine: How Federal Funding Shocks Affect University Patenting, Entrepreneurship, and Publications. *The Quarterly Journal of Economics*, 895–954.

Belenzon, S. (2012). Cumulative innovation and market value: Evidence from patent citations. *The Economic Journal*, *122*(559), 265–285.

Bercovitz, J., & Feldman, M. (2008). Academic Entrepreneurs: Organizational Change at the Individual Level. *Organization Science*, *19*(1), 69–89.

Berle, A., & Means, G. (1932). Private property and the modern corporation. *New York: Mac-millan*.

Bhaskarabhatla, A., & Hegde, D. (2014). An organizational perspective on patenting and open innovation. *Organization Science*, *25*(6), 1744–1763.

Bikard, M. (2018). Made in Academia: The Effect of Institutional Origin on Inventors' Attention to Science. *Organization Science*, *29*(5), 818–836.

Bikard, M., & Marx, M. (2019). Bridging Academia and Industry: How Geographic Hubs Connect University Science and Corporate Technology. *Management Science*, *66*(8), 3295–3798.

Birr, K. (1957). *Pioneering in industrial research: The story of the general electric research laboratory*. Public Affairs Press.

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2017). *Are ideas getting harder to find?* (Working Paper). National Bureau of Economic Research.

Braguinsky, S., & Hounshell, D. A. (2016). History and nanoeconomics in strategy and industry evolution research: Lessons from the Meiji-Era Japanese cotton spinning industry. *Strategic Management Journal*, *37*(1), 45–65.

Bucheli, M., & Wadhwani, R. D. (2014). *Organizations in time: History, theory, methods*. Oxford University Press, USA.

Cerveaux, A. (2013). Taming the microworld: DuPont and the interwar rise of fundamental industrial research. *Technology and Culture*, *54*(2), 262–288.

Clarke, H. T. (1971). Charles Edward Kenneth Mees, 1882–1960: A Biographical Memoir. *Biographical Memoirs of the National Academy of Sciences*, *1971*, 176–199.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: The two faces of R & D. *The economic journal*, *99*(397), 569–596.

Cohen, W. M., Sauermann, H., & Stephan, P. (2020). Not in the job description: The commercial activities of academic scientists and engineers. *Management Science*, *66*(9), 4108–4117.

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.

Geiger, R. L. (1986). *To advance knowledge : The growth of American research universities, 1900-1940*. New York : Oxford University Press, 1986.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning (adaptive computation and machine learning series). *Adaptive Computation and Machine Learning series*, 800.

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. (2020a). Mission-Oriented Research in a National Emergency: Lessons from the Office of Scientific Research and Development in World War II.

Gross, D. P., & Sampat, B. N. (2020b). *Organizing Crisis Innovation: Lessons from World War II*. National Bureau of Economic Research.

Gutiérrez, G., & Philippon, T. (2020). *Some facts about dominant firms*. National Bureau of Economic Research.

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.

Homburg, E. (1992). The Emergence of Research Laboratories in the Dyestuffs Industry, 1870-1900. *The British Journal for the History of Science*, *25*(1), 91–111.

Hounshell, D. (1996). The Evolution of Industrial Research in the United States. *Engines of innovation: U.S. industrial research at the end of an era*. Harvard Business School Press.

Hounshell, D., & Smith, J. K. J. (1988). *Science and corporate strategy : Du Pont R&D, 1902-1980*. Cambridge [Cambridgeshire] ; New York : Cambridge University Press, 1988.

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. *American Economic Review*, *107*(4), 1136–64.

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.

Ingram, P., Rao, H., & Silverman, B. S. (2012). History in strategy research: What, why, and how? *History and strategy* (pp. 241–273). Emerald Group Publishing Limited.

Jaffe, A. B., Lerner, J., Stern, S., & Thursby, M. C. (2007). Academic science and entrepreneurship: Dual engines of growth? *Academic Science and Entrepreneurship: Dual Engines of Growth*. Journal of Economic Behavior and Organization, 63 (4).

Jenkins, R., & Chandler, A. D. (1975). *Images and enterprise: Technology and the American photographic industry, 1839 to 1925*. Johns Hopkins University Press Baltimore.

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.

Kline, R. (1995). Construing "Technology" as "Applied Science": Public Rhetoric of Scientists and Engineers in the United States, 1880-1945. *Isis*, *86*(2), 194–221.

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., Raff, D. M., & Temin, P. (1999). *Learning by Doing in Markets, Firms, and Countries*. University of Chicago Press.

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.

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. *Lessons from the Quota Acts*.

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.

Murmann, J. P., & Landau, R. (1998). On the making of competitive advantage: The development of the chemical industries in Britain and Germany since 1850. In A. Arora, R. Landau, & N. Rosenberg (Eds.), *Chemicals and long-term economic growth: Insights from the chemical industry*.

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

Nelson, R. R. (1959). The economics of invention: A survey of the literature. *The Journal of Business*, *32*(2), 101–127.

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.

Perkmann, M., Salandra, R., Tartari, V., McKelvey, M., & Hughes, A. (2021). Academic engagement: A review of the literature 2011-2019. *Research Policy*, *50*(1), 104114.

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.

Pillai, S. D., Goldfarb, B., & Kirsch, D. (2024). Lovely and likely: Using historical methods to improve inference to the best explanation in strategy. *Strategic Management Journal*, (n/a).

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.

Riordan, M., & Hoddeson, L. (1997). *Crystal fire: The invention of the transistor and the birth of the information age*. WW Norton & Company.

Rosenberg, N. (1990). Why do firms do basic research (with their own money)? *Research policy*, *19*(2), 165–174.

Rothaermel, F. T., Agung, S. D., & Jiang, L. (2007). University entrepreneurship: A taxonomy of the literature. *Industrial and corporate change*, *16*(4), 691–791.

Rowland, H. A. (1883). A plea for pure science. *Science*, *2*(29), 242–250.

Sauermann, H., & Cohen, W. M. (2010). What makes them tick? Employee motives and firm innovation. *Management Science*, *56*(12), 2134–2153.

Service, N. R. C. (R. I. (1931). *Industrial Research Laboratories of the United States, Including Consulting Research Laboratories*. RR Bowker Company.

Servos, J. W. (1994). Changing partners: The Mellon Institute, private industry, and the federal patron. *Technology and Culture*, *35*(2), 221–257.

Shvadron, D. (2024). *Bread Upon the Waters: Corporate Science and the Benefits from Follow-on Public Research*.

Silverman, B. S., & Ingram, P. (2017). Asset ownership and incentives in early shareholder capitalism: Liverpool shipping in the eighteenth century. *Strategic Management Journal*, *38*(4), 854–875.

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

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

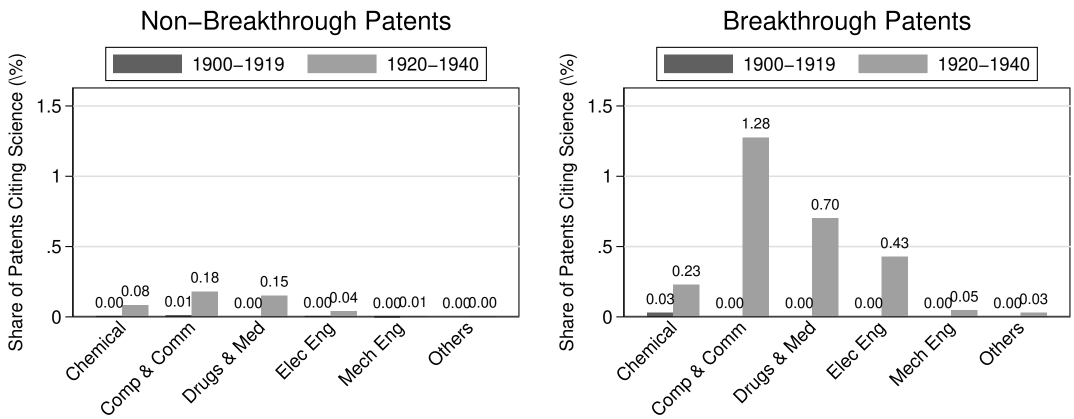
Terman, F. E. (1976). A brief history of electrical engineering education. *Proceedings of the IEEE*, *64*(9), 1399–1407.

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.

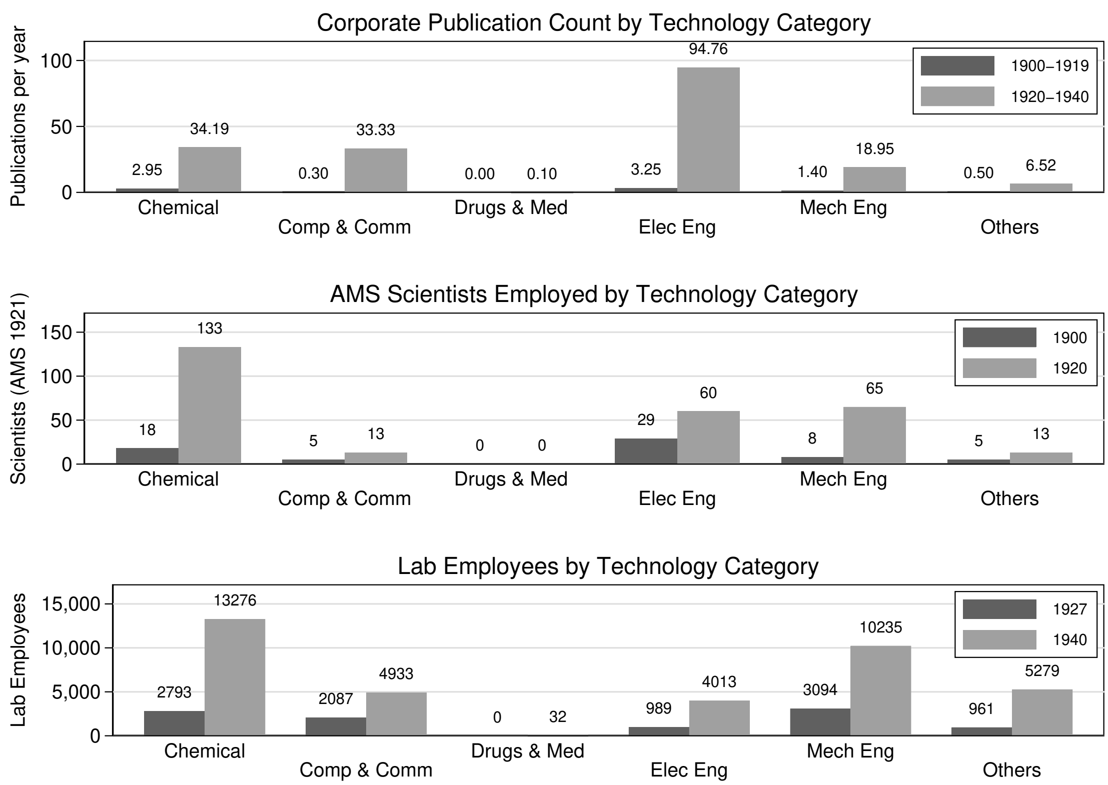
Zucker, L. G., Darby, M. R., & Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management science*, *48*(1), 138–153.

Figure 1: PATENT CITATIONS TO SCIENCE, BY BREAKTHROUGH PATENT STATUS (KPST) AND TECHNOLOGY CATEGORY



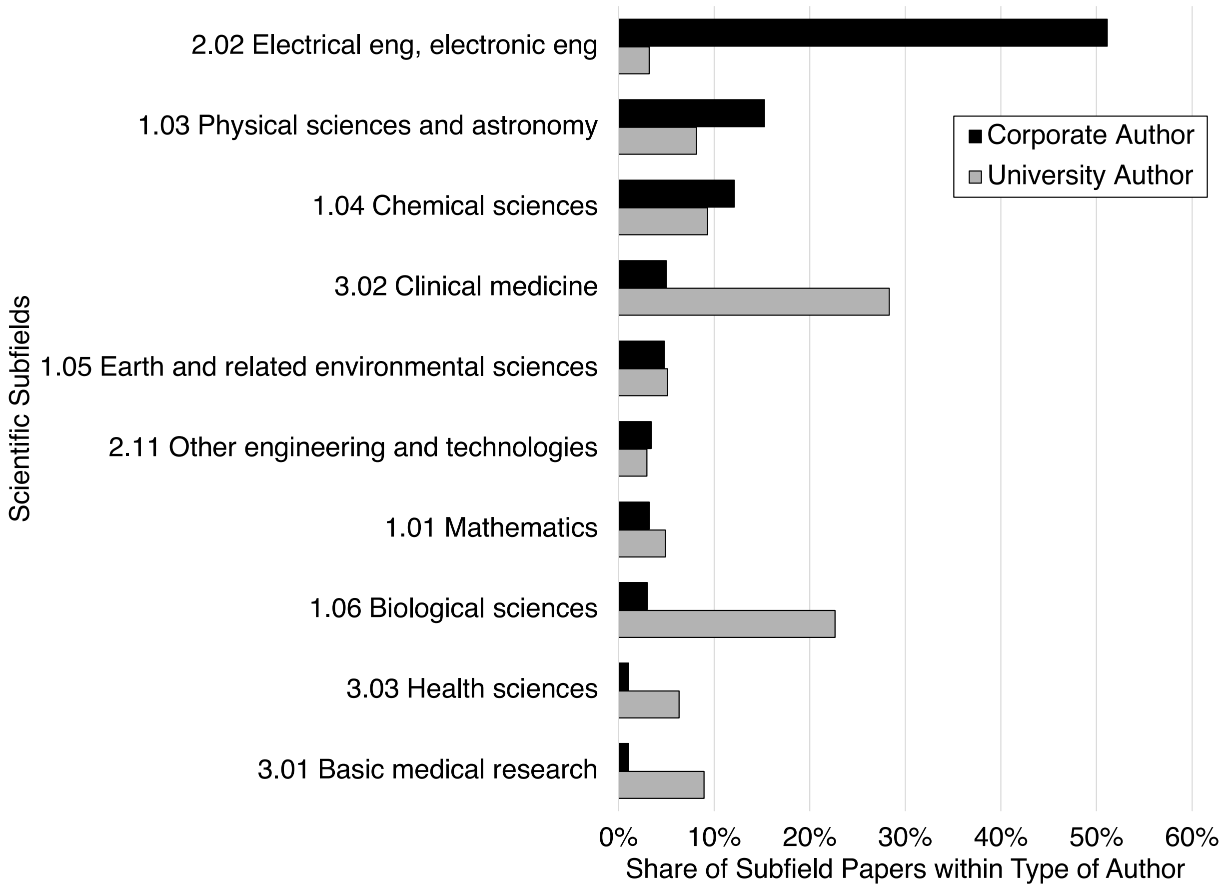
Notes: The right panel plots the percentage share of "breakthrough" patents that cite scientific articles for patents issued between 1900 and 1919 and 1920 and 1940, by NBER categories. The left panel plots the same for all other ("non-breakthrough") patents. Breakthrough patents are defined as those that are in the top 10% of the importance measures from Kelly et al. (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



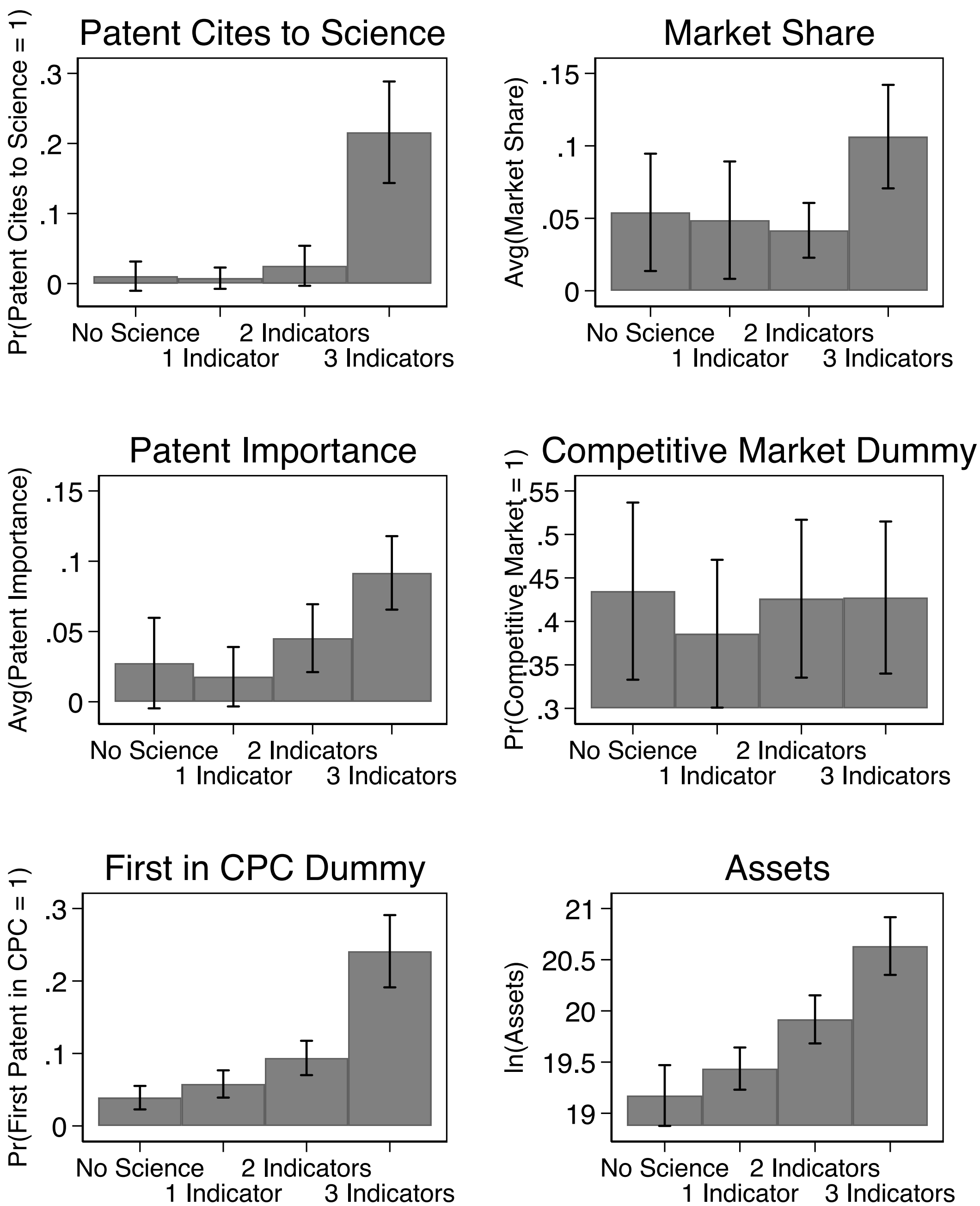
Notes: The top panel plots total corporate publications in each NBER patent category per year produced between 1900 and 1919 and between 1920 and 1940 respectively. The middle panel plots the number of scientists from the AMS directory affiliated with firms in each category in 1900 and 1920. The bottom 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: SCIENTIFIC PUBLICATIONS BY SUBFIELD AND TYPE OF AUTHOR



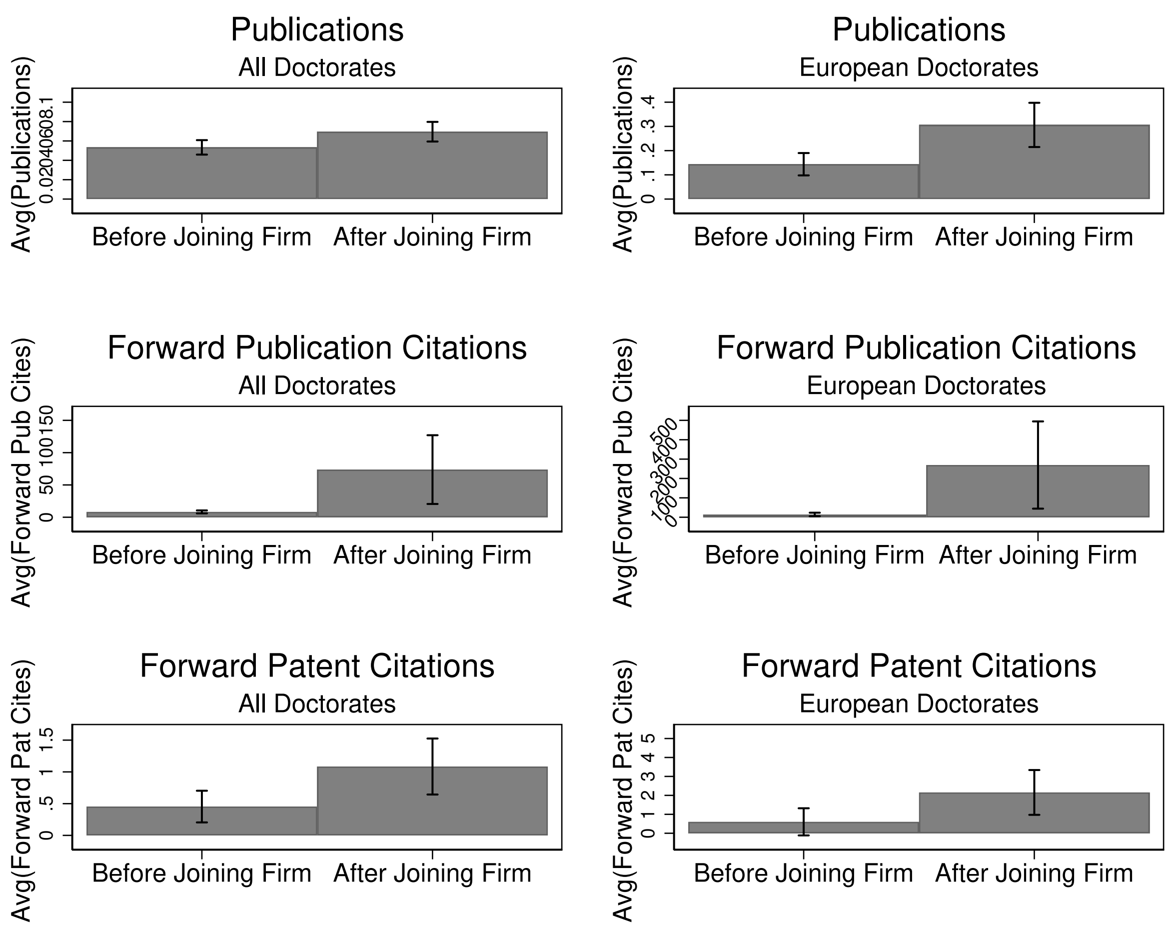
Notes: The black bars divide the number of corporate papers in a scientific subfield by all corporate papers. The gray bars divide the number of U.S. university papers in a scientific subfield by all U.S. university papers. Scientific subfields with at least 1,000 papers during the sample period are included in the analysis.

Figure 4: 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.

Figure 5: PUBLICATION CHARACTERISTICS OF CORPORATE SCIENTISTS WITH DOCTORATES



Notes: The top-left panel plots average MAG publications 10 years before ("Before Joining Firm") and 10 years after ("After Joining Firm") by all doctorate holders with a corporate affiliation in the 1938 edition of the American Men of Science (AMS) Directory (N=747). The top-right panel plots this number for a subset of corporate scientists that have received their doctoral degrees from a European institution (N=50). The mid and bottom panels respectively plot the average publication and patent citations received by the papers in the top panel.

Table 1: SCIENTIFIC PAPER PROXIMITY TO U.S. PATENTS (LINEAR PROBABILITY MODEL)

Dependent Variable	Cited by Patent		Within 10 Closest	
	(1)	(2)	(3)	(4)
Corporate Publication Dummy	0.014	0.007	0.310	0.042
	(0.002)	(0.003)	(0.005)	(0.007)
Journal Impact Factor	0.000	-0.002	-0.229	-0.014
	(0.001)	(0.002)	(0.011)	(0.020)
Average of Dep Var	0.003	0.003	0.646	0.646
Year Dummies	Yes	Yes	Yes	Yes
Subfield Dummies	No	Yes	No	Yes
Journal Dummies	No	Yes	No	Yes
R^2	0.006	0.020	0.062	0.221
N	38,887	38,816	38,887	38,816

Notes: The unit of analysis is a scientific paper. The sample is limited to American science or engineering papers published from 1926 to 1940 in Microsoft Academic Graph (MAG). "Corporate Publication Dummy" is equal to one if the paper was published by one of the sample firms described in Section 3, and zero if it is authored by a U.S. university. The dependent variable for Columns 1 and 2 is equal to one if the paper has been cited by a patent; for Columns 3 and 4 it is equal to one if the paper is ranked within the 10 most textually similar papers to a patent. Textual similarity is calculated through the SPECTER algorithm (details in Arora et al. (2023)). All specifications include fixed effects for paper publication year. Columns 2 and 4 add fixed effects for scientific subfield and journal.

Table 2: CROSS TABULATION OF MEASURES OF CORPORATE SCIENCE

	Does Not Publish	Publishes	Publishing Firm Share
Does Not Operate Lab	116	42	27%
Operates Lab	149	159	52%
	Does Not Employ AMS	**Employs AMS**	**AMS Firm Share**
Does Not Operate Lab	121	37	23%
Operates Lab	114	194	63%
	Does Not Employ AMS	**Employs AMS**	**AMS Firm Share**
Does Not Publish	174	91	34%
Publishes	61	149	71%

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 1938 edition of AMS.

Table 3: 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 4: CORPORATE SCIENTIFIC PAPERS AND TECHNOLOGICAL & MARKET LEADERSHIP (POISSON)

Dependent Variable	Publication Count							
	Technological Leadership				Market Leadership			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for Patent Cite to Science	2.135			1.107				
	(0.446)			(0.468)				
Dummy for First Patent in CPC		1.770		0.714				
		(0.247)		(0.144)				
Patent Importance			4.327	2.874				
			(0.750)	(0.855)				
ln(Assets, 1926-1930)					1.046			0.684
					(0.283)			(0.250)
Market Share						6.337		1.834
						(0.687)		(1.536)
Dummy for Competitive Market							-1.045	-0.977
							(0.503)	(0.508)
ln(Assets)	0.617	0.667	0.525	0.479		0.494	0.778	
	(0.205)	(0.220)	(0.140)	(0.129)		(0.084)	(0.105)	
Distance to Universities	0.016	-0.010	0.029	0.057	-0.052	-0.050	-0.045	-0.031
	(0.031)	(0.027)	(0.035)	(0.031)	(0.042)	(0.062)	(0.056)	(0.045)
PhD Graduates	0.014	0.012	0.696	0.348	0.051	0.041	0.130	0.113
	(0.018)	(0.019)	(0.133)	(0.133)	(0.022)	(0.019)	(0.093)	(0.062)
Average of Dep Var	0.688	0.688	0.939	0.939	0.628	0.922	0.624	0.716
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Pseudo-R^2	0.670	0.649	0.716	0.733	0.605	0.728	0.372	0.430
Number of Firms	422	422	397	397	355	273	417	272
Number of Obs	3,855	3,855	2,640	2,640	4,860	2,747	4,252	4,080

Notes: The analysis is at the firm-year level. Columns 1-4 regress corporate publications against firm-level measures of technologcial leadership; Columns 5-8 regress them against measures of market leadership. "ln(Assets)" takes the natural log of concurrent assets; "ln(Assets, 1926-1930)" takes the natural log of average assets between 1926 and 1930 (all dollar amounts in this paper are deflated to 2005 using https://www.measuringworth.com/datasets/usgdp12/result.php). "Market Share" averages firm market share at the 3-digit industry-level between 1926 and 1930. See Appendix Table B1 for further details on variable construction. Standard errors are clustered at the firm level. Standard errors are clustered at the firm level.

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

Dependent Variable	Publication Count							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap	26.343	41.509	-1.338	-0.465	16.732	-133.178	3.041	34.389
	(7.624)	(11.894)	(7.312)	(9.123)	(8.068)	(145.425)	(7.812)	(5.413)
Univ Gap × Distance to Universities		-2.311						
		(1.439)						
Univ Gap × Dummy for Patent Cite to Science			86.697					
			(14.577)					
Univ Gap × Dummy for First Patent in CPC				54.643				
				(11.620)				
Univ Gap × Patent Importance					135.141			
					(25.226)			
Univ Gap × ln(Assets, 1926-1930)						7.669		
						(6.914)		
Univ Gap × Market Share							104.177	
							(49.577)	
Univ Gap × Dummy for Competitive Market								-32.105
								(7.744)
Dummy for Patent Cite to Science			-59.299					
			(10.172)					
Dummy for First Patent in CPC				-37.103				
				(8.135)				
Patent Importance					-91.569			
					(17.488)			
ln(Assets, 1926-1930)						-4.584		
						(4.843)		
Market Share							-68.427	
							(35.518)	
Dummy for Competitive Market								21.731
								(5.495)
ln(Assets)	0.825	0.830	0.577	0.633	0.679		0.540	0.694
	(0.205)	(0.205)	(0.101)	(0.128)	(0.097)		(0.063)	(0.070)
Distance to Universities	-0.036	1.626	0.043	0.025	-0.001	-0.049	-0.062	-0.079
	(0.040)	(1.014)	(0.026)	(0.027)	(0.042)	(0.037)	(0.064)	(0.060)
PhD Graduates	0.072	0.069	0.008	0.012	0.466	0.065	0.040	0.133
	(0.028)	(0.029)	(0.014)	(0.017)	(0.096)	(0.024)	(0.021)	(0.066)
Average of Dependent Variable	0.688	0.688	0.688	0.688	0.939	0.628	0.922	0.624
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.653	0.655	0.734	0.702	0.743	0.655	0.744	0.458
Number of Firms	422	422	422	422	397	355	273	417
Number of Observations	3,855	3,855	3,855	3,855	2,640	4,860	2,747	4,252

Notes: The analysis is at the firm-year level. "ln(Assets)" takes the natural log of concurrent assets; "ln(Assets, 1926-1930)" takes the natural log of 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. Controls for the level of European and American publications are included. See Appendix Table B1 for further details on variable construction. Standard errors are clustered at the firm level.

Table 6: CORPORATE SCIENCE AND FIRM PERFORMANCE

Dependent Variable	Top 5% Market Value (KPSS)	Top 5% Forward Cites	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)	(9)
Sample	All	All	All	Univ Gap Below Mean	Univ Gap Above Mean	No	Yes	Univ Gap Below Mean	Univ Gap Above Mean
ln(Publication Stock$_{t-1}$)	0.313	0.187	0.165	0.123	0.206	0.110	0.204	0.235	0.325
	(0.118)	(0.100)	(0.055)	(0.130)	(0.059)	(0.091)	(0.081)	(0.148)	(0.094)
ln(Patent Stock$_{t-1}$)	0.165	0.703	0.126	0.193	0.098	0.144	0.119	0.197	-0.004
	(0.095)	(0.086)	(0.036)	(0.057)	(0.038)	(0.051)	(0.077)	(0.127)	(0.097)
ln(Assets$_{t-1}$)	0.927	-0.103	0.748	0.733	0.739	0.736	0.778	0.701	0.927
	(0.185)	(0.087)	(0.050)	(0.077)	(0.063)	(0.073)	(0.102)	(0.167)	(0.084)
Distance to Universities	-0.092	-0.026	0.007	0.013	-0.007	0.017	-0.040	-0.061	-0.032
	(0.067)	(0.026)	(0.012)	(0.033)	(0.015)	(0.013)	(0.027)	(0.036)	(0.035)
PhD Graduates$_{t-1}$	0.092	0.065	0.001	-0.000	0.003	-0.003	-0.006	-0.011	-0.003
	(0.040)	(0.015)	(0.003)	(0.005)	(0.004)	(0.006)	(0.009)	(0.014)	(0.010)
Average of Dependent Variable	1.574	0.844	19.418	19.348	19.490	18.911	20.006	19.951	20.137
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2			0.731	0.717	0.806	0.688	0.773	0.778	0.876
Number of Firms	321	418	321	167	154	219	139	76	63
Number of Observations	1,888	3,916	3,333	1,681	1,651	659	274	146	121

Notes: The analysis is at the firm-year level. Columns 1 and 2 regress counts of valuable patents against corporate investments in science (corporate papers) using a Poisson regression. Columns 3-9 regress log of market capitalization against corporate investment in science using OLS. Columns 4 and 5 split the sample by mean values of the "University Gap" measure based on European and American publications. Columns 6 and 7 split the sample by whether the firm both employs lab personnel as well as AMS scientists. Columns 8 and 9 split the Column 7 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.

Table 7: Average Pairwise Jaccard Similarity Scores for Scientists' Topics

	Corporate Scientist	Non-Corporate: U.S. PhD	Non-Corporate: European PhD
Corporate Scientist	2.719 (0.051) N=16,984		
Non-Corporate: U.S. PhD	2.478 (0.016) N=149,937	2.676 (0.008) N=546,524	
Non-Corporate: European PhD	2.190 (0.024) N=51,071	2.379 (0.010) N=324,148	2.328 (0.022) N=52,691
Number of Scientists	356	3,940	1,200

Notes: The sample is limited to 5,496 scientists with doctorates in the 1921 American Men of Science (AMS). Six sets of scientist pairs (in the same field) are analyzed: the first column shows summary statistics of Jaccard similarities (ranging from 0 to 1) of topics listed in AMS for 356 corporate scientists 1) against themselves in the first row; 2) against scientists without corporate affiliations that hold U.S. doctorates in the second row; 3) against non-corporate scientists that hold European doctorates. The second column limits the sample to 3,940 non-corporate scientists with U.S. PhDs and calculates similarities within this set in the second row, and against non-corporate scientists with European PhD in the third row. The third column limits the sample to non-corporate scientists with European PhDs. Each quadrant calculates the mean and standard errors (in parentheses) of pairwise Jaccard similarities of the constituent scientist pairs (multiplied by 100 to aid with interpretation), followed by corresponding number of pairs. A quadrant is shaded darker when the mean pairwise Jaccard similarity score is higher.

For Online Publication

Appendix A Analytical Model

Section 2 that recounts the rise of American corporate science 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.

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.[50]

A.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

[50] 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.

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 $\tilde{\phi(u)}d_1$. As noted, innovations may be based on external discoveries and inventions. Thus, we assume that $\tilde{\phi(u)}$ 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 $v_0=kd_0-\frac{c_{00}}{2}d_0^2-bd_1-\frac{c_{11}}{2}d_1^2+c_{01}d_1d_0-\phi(r_0,\lambda)d_0-\frac{\gamma}{2}r_0^2$.

A.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.

A.2.1 Focal Firm Research and Innovation

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

$$\begin{aligned}\frac{\partial d_0}{\partial r_0}&=\frac{c_{00}}{D}(-\frac{\partial\phi}{\partial r_0})\\ \frac{\partial d_1}{\partial r_0}&=\frac{c_{01}}{D}(-\frac{\partial\phi}{\partial r_0})\end{aligned} \tag{A1}$$

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.

A.2.2 Public Science and Innovation

The response of innovation output to public science is

$$\begin{aligned}\frac{\partial d_0}{\partial u}&=\frac{-1}{D}\left(c_{00}\frac{\partial\phi}{\partial u}+c_{01}\frac{\partial\tilde{\phi}}{\partial u}\right)\\ \frac{\partial d_1}{\partial u}&=\frac{-1}{D}\left(c_{00}\frac{\partial\tilde{\phi}}{\partial u}+c_{01}\frac{\partial\phi}{\partial u}\right)\end{aligned} \tag{A2}$$

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

$$\frac{\partial d_0}{\partial u}+\frac{\partial d_1}{\partial u}=\frac{-1}{D}\left(\frac{\partial \phi}{\partial u}+\frac{\partial \tilde{\phi}}{\partial u}\right)(c_{00}+c_{01})\geq 0 \tag{A3}$$

This implies that innovation on average increases with public science.

A.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

$$-\frac{\partial \phi}{\partial r_0}d_0+\frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial r_0}=\gamma r_0 \tag{A4}$$

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 A1 and gathering terms, Equation A4 can be rewritten as

$$-\frac{\partial \phi}{\partial r_0}\left(\frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}+d_0\right)=\gamma r_0 \tag{A5}$$

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, $c_{01}\leq 0$. [51]

A.4 Innovation Leadership

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

$$\begin{aligned}\frac{\partial v}{\partial k}&=d_0+\frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial k}\\&=d_0+\frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}>0 \text{ at an interior maximum}\\\frac{\partial \tilde{v}}{\partial k}&=\frac{\partial \tilde{\Pi}}{\partial d_0}\frac{\partial d_0}{\partial k}=\frac{\partial \tilde{\Pi}}{\partial d_0}\frac{c_{00}}{D}<0\end{aligned} \tag{A6}$$

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.

$$\begin{aligned}\frac{\partial^2 v}{\partial k\partial r_0}&=\frac{\partial d_0}{\partial r_0}+\frac{c_{01}}{D}(-c_{11}\frac{\partial d_1}{\partial r_0}+c_{00}\frac{\partial d_0}{\partial r_0})\\&=(-\frac{\partial \phi}{\partial r_0})\left(\frac{c_{00}}{D}+\frac{c_{01}^2}{D}(c_{00}-c_{11}\right)>0\\\frac{\partial^2 \tilde{v}}{\partial k\partial r_1}&=\frac{c_{00}}{D}(-c_{11}\frac{\partial d_0}{\partial r_0}+c_{01}\frac{\partial d_1}{\partial r_1})=\frac{c_{00}}{D}(-\frac{\partial \tilde{\phi}}{\partial r_1})c_{01}(c_{00}-c_{11})\leq 0 \iff c_{01}\leq 0\end{aligned} \tag{A7}$$

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

This result points to why the follower may not invest in research. Equation A7 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.

A.5 Public Science

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

A.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

$$\begin{aligned}\frac{\partial v}{\partial u} &= -d_0 \frac{\partial \phi}{\partial u} + \frac{\partial \Pi}{\partial d_1}\frac{\partial d_1}{\partial u} \\ &= -\frac{\partial \phi}{\partial u}\left(d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}\right) - c_{00}\frac{\partial \Pi}{\partial d_1}\frac{\partial \tilde{\phi}}{\partial u}\end{aligned} \tag{A8}$$

Although the first term is positive if the firm invests in internal research, by Equation A4, 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.

A.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.

$$\begin{aligned}\frac{\partial^2 v}{\partial r_0 \partial u} &= \left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} + d_0\left(-\frac{\partial^2 \phi}{\partial r_0 \partial u}\right) + \frac{\partial \Pi}{\partial d_1}\frac{\partial^2 d_1}{\partial r_0 \partial u} + \frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u} \\ &\quad \text{substituting and collecting terms} \\ &= \left(-\frac{\partial \phi}{\partial r_0}\right)\frac{\partial d_0}{\partial u} - \frac{\partial^2 \phi}{\partial r_0 \partial u}\left(d_0 + \frac{\partial \Pi}{\partial d_1}\frac{c_{01}}{D}\right) + \frac{\partial d_1}{\partial r_0}\frac{\partial^2 \Pi}{\partial d_1 \partial u}\end{aligned} \tag{A9}$$

The first term in Equation A9 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 A9 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, A9 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

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

$$= \frac{\partial d_1}{\partial r_0}\frac{1}{D}\left[-c_{11}c_{00}(-\frac{\partial\tilde{\phi}}{\partial u}) - c_{11}c_{01}(-\frac{\partial\phi}{\partial u}) + c_{01}c_{00}(-\frac{\partial\phi}{\partial u}) + c_{01}^2(-\frac{\partial\tilde{\phi}}{\partial u})\right] \tag{A10}$$

collecting terms and substituting

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

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 A10 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.

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

Recall from Equation A7 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

$$\frac{\partial r_0}{\partial k} = -\frac{\partial\phi}{\partial r_0}\left(\frac{c_{00}}{D} + \frac{c_{01}^2}{D}(c_{00} - c_{11})\right)\left(-\frac{\partial^2 v}{\partial r_0^2}\right)^{-1} > 0 \tag{A11}$$

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 $\left(-\dfrac{\partial^2 v}{\partial r_0^2}\right)^{-1}$ is constant. Then, $\dfrac{\partial^2 r_0}{\partial k \partial u} \geq 0 \iff -\dfrac{\partial^2 \phi}{\partial r_0 \partial u} \geq 0$.

Table A1: 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 { Patent Cites to Science First Patent in CPC Patent Importance Market Share	+	+
Univ Science	$\frac{\partial r_0}{\partial u}$	u { Eur. Scientific Pubs Cites to Eur. Journals	+ if Strategic Substitutes	Ambiguous
Proximity to Frontier and Univ Science	$\frac{\partial^2 r_0}{\partial k \partial u}$	Same As Above	+	-

Appendix B Data Appendix

Table B1: 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 (1938)
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 (in 100 miles)	Wilson (1932)
PhD Graduates	Number of science & engineering doctoral dissertations relevant to a firm's patents (divided by 100)	Arora, Belenzon, Cioaca, Sheer, and Zhang (2023)

Notes: All dollar amounts (market value, sales, and assets) are deflated to 2005 using https://www.measuringworth.com/datasets/usgdp12/result.php.

Table B2: SUMMARY STATISTICS OF MAIN VARIABLES

	Count	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	8.56	7.90	3.07	6.83	42.44
PhD Graduates	6990	13.52	16.35	12.54	0.00	37.12
Lab Size	2330	44.61	0.00	220.63	0.00	4669.00
Publications, Annual Count	6990	0.47	0.00	3.73	0.00	88.00
AMS Scientists	6058	3.07	0.00	17.45	0.00	379.00
Patents, Annual Count	6990	13.21	1.00	55.04	0.00	838.00
Patents Within Top 5% Xi (KPSS)	6990	0.53	0.00	4.78	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

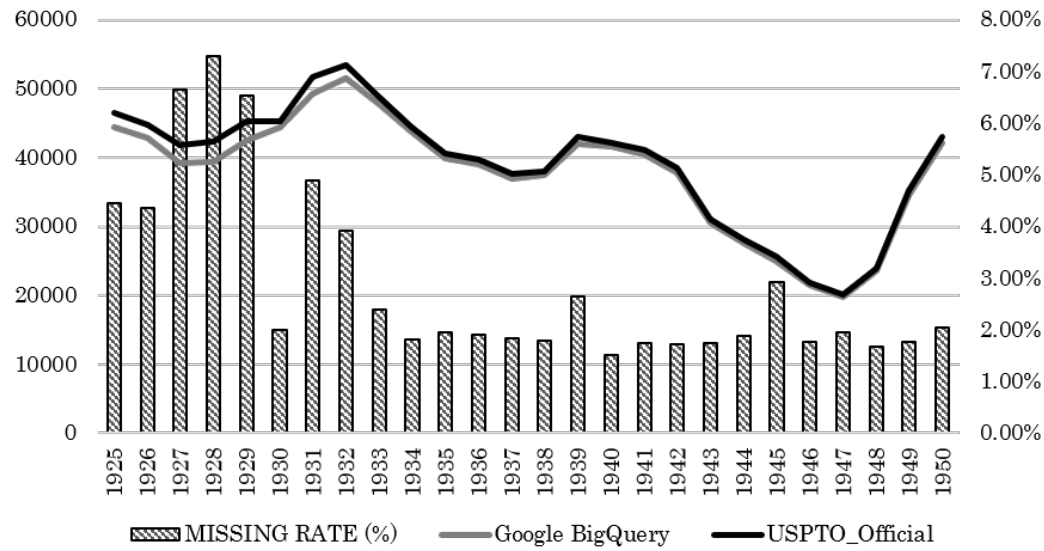
Notes: Observations are at the firm-year level. Distance to Universities is in hundreds of miles. Number of PhD Graduates is also divided by 100. The sample period is between 1926 and 1940. All dollar amounts are deflated to 2005 dollars using https://www.measuringworth.com/datasets/usgdp12/result.php. Details on variable definition and data sources can be found in Appendix Table B1.

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

B.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 B1 to ensure that our data source does not have coverage issues.[52] We find that the missing rate is around 3.43%; there are an average of 42,476 utility patents granted per year from 1926 to 1940.

Figure B1: 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 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.[53] Moreover, we manually check the names of 620 patentees with above 100 patents

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

[53] 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.

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).[54]

B.1.2 Matching Corporations to Publications

Table B3: 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 from 1900 to 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.[55] 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 firm. We also cross-tabulate 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.

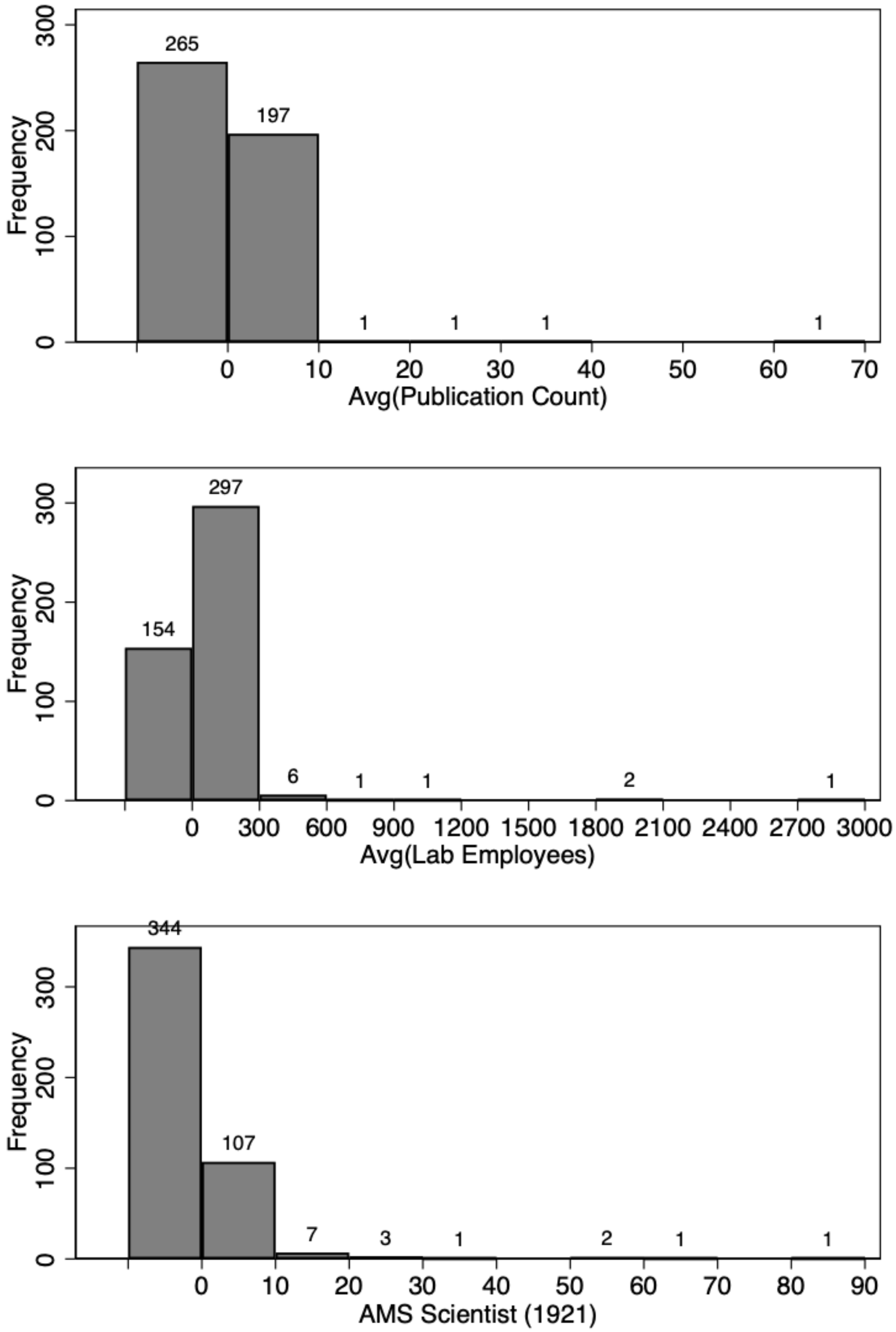
B.1.3 Matching Corporations to Industrial Research Laboratories

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 B3) vs 2 pages for DuPont (figure B4)) and the fields are

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

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

Figure B2: 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 from 1926 to 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.

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 who 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 B3: 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 B4: 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 B4: 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 from 1926 to 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.

B.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). Out of a total of 9,557 scientists, we identify 653 that work for 122 firms in our sample. To increase coverage of our sample period, we also search through around 30,000 scientists in the 1938 edition of AMS and find 3,322 working at 231 sample firms.

Matching AMS Scientists to Scientific Publications — To reduce instances of false positive matches, we exact-match author names from Microsoft Academic Graph (MAG) against the full first and last names of scientists transcribed from the 1938 edition of AMS. We find 1,259 authors publishing 3,564 MAG papers between 1900 and 1940. Figure B6 plots corporate papers that make citations to Europe[56] (normalized by all of the focal firm's papers) against European PhDs employed at firms (normalized by all AMS scientists working at the focal firm) and finds a positive correlation (r=0.245).

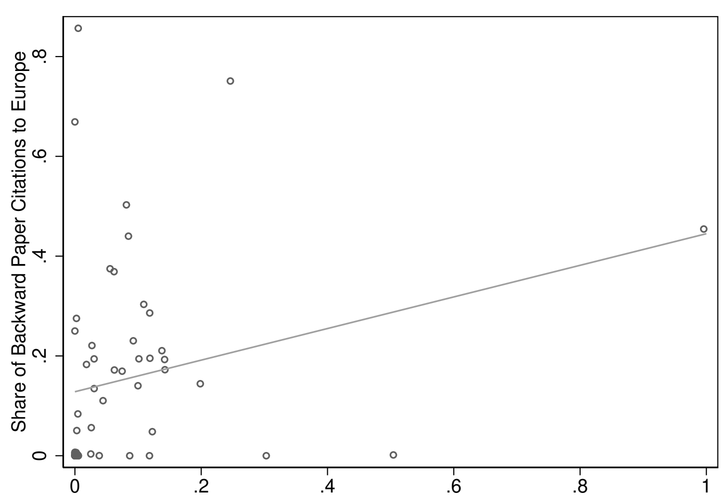
[56] Since MAG does not have location information for the publisher, we link MAG journals to Web of Science to identify European journals (further details are discussed in Appendix Section C).

Figure B5: 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 (AMS) Directory.

Figure B6: CITATIONS TO EUROPE VS EMPLOYMENT OF EUROPEAN PHDS



Notes: Level of analysis is at the firm level. Y-axis refers to the share of European papers found in the backward paper citations made by corporate publications between 1900 and 1940. X-axis refers to the share of scientists with European PhDs out of total PhDs employed as indicated in the 1938 edition of the American Men of Science Directory.

Calculating Topic Similarity between Scientist Pairs — We use the 1921 edition of AMS data from Moser et al. (2022) to compare research topics of scientists. We remove all stopwords including prepositions ("of") and conjunctions ("and") and generate pairs of scientists working in the same field. In the case of Gilbert Lewis in Figure B5, the scientist's field is "Chemistry," while his topics are "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".

B.2 Measuring University Human Capital Relevant to Corporations

B.2.1 Firm Geographical Distance to Universities

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 B7, 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[57] and extract the FIPS (Federal Information 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.

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).[58] 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.

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".[59]

B.2.2 Identifying Dissertations Relevant to Firms

We also measure human capital relevant to corporate innovations by counting the number of PhD dissertations from the ProQuest Dissertations and Theses (PQDT) database close to firm patents. PQDT archives over five million dissertations submitted in fulfillment of graduate degrees in the United States from 1637

[57] Available from https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BPC15W
[58] Available form https://babel.hathitrust.org/cgi/pt?id=uc1.b3509036&view=1up&seq=7).
[59] Package download link: urlhttps://ideas.repec.org/c/boc/bocode/s457147.html

to present day. We focus on a subset of 27,766 dissertations in the natural sciences[60] submitted to degree granting institutions in the United States from 1925 to 1941.

Since PhD dissertations are not typically cited by patents or publications, we assess the relevance of each thesis to a firm's innovations based on the textual similarities between theses and patents using a deep learning algorithm, SPECTER. SPECTER (Scientific Paper Embeddings using Citation-informed TransformERs) is a specialized Bidirectional Encoder Representations from Transformers (BERT) model trained on 146,000 scientific papers (containing 26.7 million words) from Semantic Scholar and their forward citations. SPECTER has been compared to several other deep learning models specialized in technical documents, including BERT, PatentBERT, BERT for Patents, PatentSBERTa, SciBERT, RoBERTa, and ALBERT. For detailed descriptions of how SPECTER is implemented, please see Arora et al. (2023).

For each corporate patent, we identified the top 1,000 most similar dissertations awarded in years $[t-1, t+1]$, where t is the patent grant year. Because a PhD dissertation could be similar to multiple patents, we used the maximum similarity score between a dissertation and all patents it was found similar to. The "PhD Graduates" variable was then constructed by summing over the maximum similarity scores at the firm-year level. Hence, the measure is a weighted sum of PhD dissertations where the weights are the (maximal) similarity scores (ranging from zero to one) between American science & engineering dissertations and patents owned by the focal firm.

B.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.[61] 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.[62] For each ultimate owner-year, we calculate the share of each 3-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.[63]

B.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

[60] We use Proquest's subject thesaurus and classify as natural science the following categories: "Pure sciences". "Biological sciences", "Health and environmental sciences", "Applied sciences", "Earth science". This excludes categories in the humanities and the social sciences such as "Language", "Literature", and "Psychology." The ProQuest Thesaurus of subject terms is a proprietary vocabulary that has been developed over many years by ProQuest Taxonomy experts. The thesaurus is actively managed to ensure that terms remain accurate and up to date, while new terms are added regularly to ensure that our indexing remains relevant in capturing new and emerging concepts and topics.

[61] 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.

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

[63] 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)

cases of public utility apex companies controlling industrial companies).

B.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).

B.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

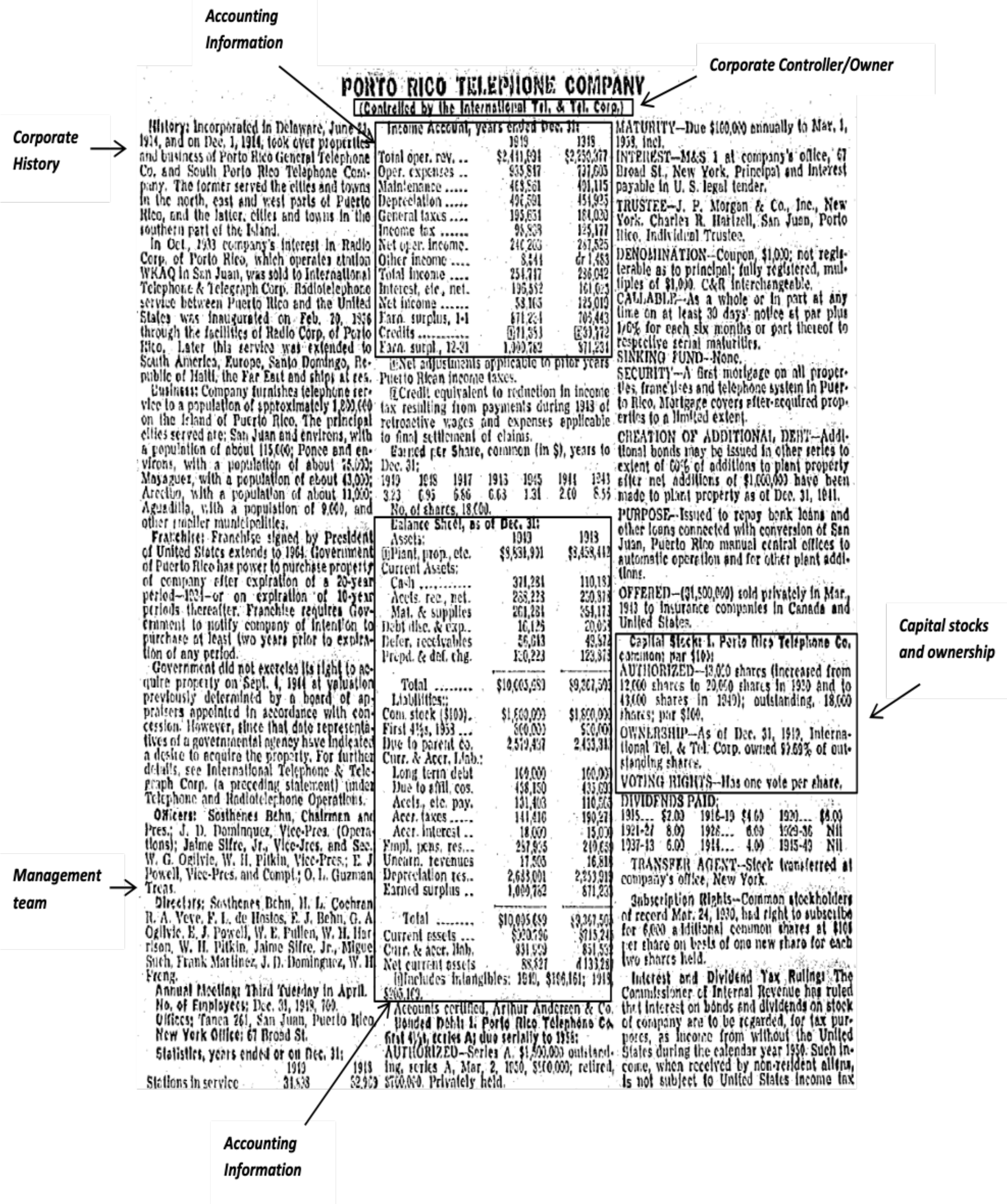
- Securities and Exchange Commission (SEC) Annual Reports: www.sec.gov/about/annrep.shtml

- 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)

Figure B7: 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 1,200,000 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. 1, 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. Fullen, 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:

	1919	1918
Stations in service	31,838	52,929

Income Account, years ended Dec. 31:

	1919	1918
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	195,631	184,030
Income tax	95,838	125,177
Net oper. income.	240,203	267,525
Other income	8,541	dr 1,483
Total income	254,747	236,042
Interest, etc., net.	196,582	161,023
Net income	58,165	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:

1919	1918	1917	1913	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
(i)Plant, prop., etc.	$9,831,901	$8,458,412
Current Assets:		
Cash	371,281	110,181
Accts. rec., net.	268,223	250,818
Mat. & supplies	281,281	354,173
Debt disc. & exp..	16,125	20,053
Defer. receivables	56,013	49,574
Prepd. & def. chg.	130,223	123,873
Total	$10,005,689	$9,367,503
Liabilities:		
Com. stock ($100).	$1,800,000	$1,800,000
First 4½s, 1958 ...	900,000	500,000
Due to parent co.	2,519,437	2,433,311
Curr. & Accr. Liab.:		
Long term debt	100,000	100,000
Due to affil. cos.	438,150	435,690
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..	2,685,001	2,353,919
Earned surplus ..	1,000,762	871,231
Total	$10,005,689	$9,367,503
Current assets ...	$920,796	$715,246
Curr. & accr. liab.	831,959	851,533
Net current assets	88,827	d 133,287

(i)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, $900,000; retired, $600,000. 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 1930 and to 43,000 shares in 1949); outstanding, 18,000 shares; par $100.

OWNERSHIP—As of Dec. 31, 1949, International Tel. & Tel. Corp. owned 97.69% 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).

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

Appendix C Gaps in University Science

C.1 Details on Calculating Gaps in University Science

Figure C1 provides an example of how the publications-based gap calculation is done for Black & Decker Company, which patents in four patent classes from 1926 to 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 from 1947 to 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:

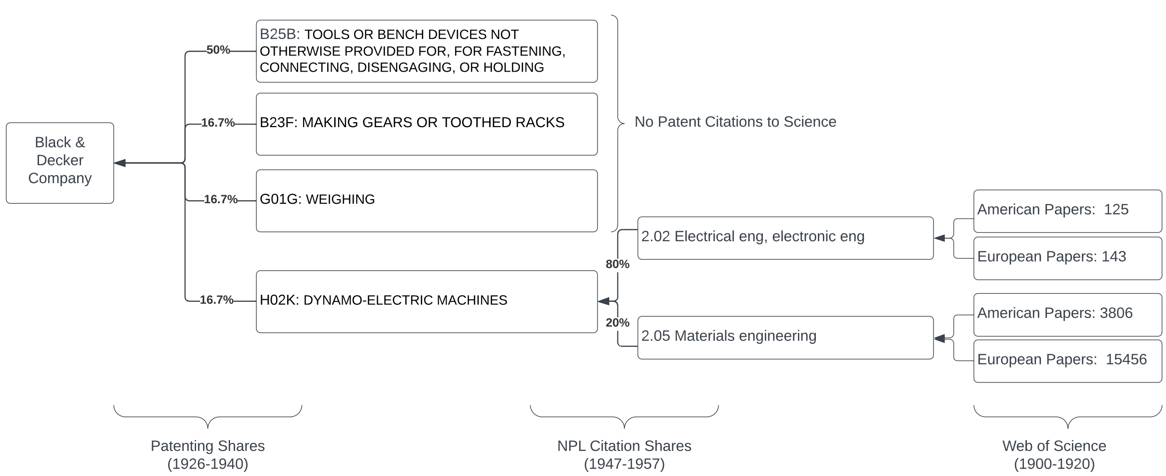
$$Eur.\ Papers_{H02K,EE} + Eur.\ Papers_{H02K,ME} = 0.8 \times 143 + 0.2 \times 15456 = 3205.6 \tag{C1}$$

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 $Eur.\ Papers_{B\&D,H02K} + Eur.\ Papers_{B\&D,G01G} + Eur.\ Papers_{B\&D,B23F} + Eur.\ Papers_{B\&D,B25B} = .167 \times 3205.6 + .167 \times 0 + .167 \times 0 + .5 \times 0 = 538.5$.

The American papers relevant to the firm are similarly calculated as $US\ Papers_{B\&D,H02K} + US\ Papers_{B\&D,xG01G} + US\ Papers_{B\&D,B23F} + US\ Papers_{B\&D,B25B} = .167 \times 862 + .167 \times 0 + .167 \times 0 + .5 \times 0 = 144$. It follows that *Gap in university science, 1900-20* value for Black & Decker is $538.5/(538.5 + 144) = .79$.

For the period from 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).

Figure C1: PUBLICATION-BASED SCIENTIFIC GAP CALCULATION FOR BLACK & DECKER



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

Journal Citation-Based Gap — 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.[64] 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.

For articles published from 1900 to 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.

C.2 Comparison Between Gap Measures

Table C1: 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

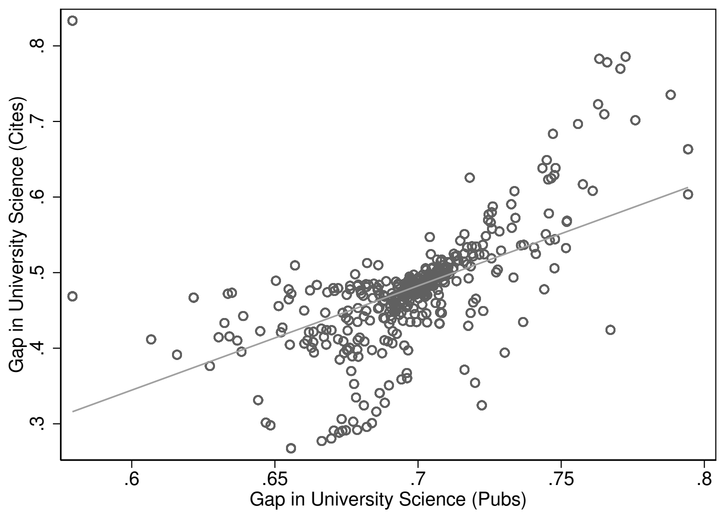
Notes: This table presents the number of citation-weighted articles (from WoS) that have non-missing subject and affiliation fields. The "Ratio" column for the Publications sub-columns divides the number of European-affiliated papers (published globally) divided by American-affiliated papers. The rows 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 C1 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

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

1,306 total citations).[65] Excluding this outlier, the correlation 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 C2).

Figure C2: COMPARISON OF GAPS IN UNIVERSITY SCIENCE

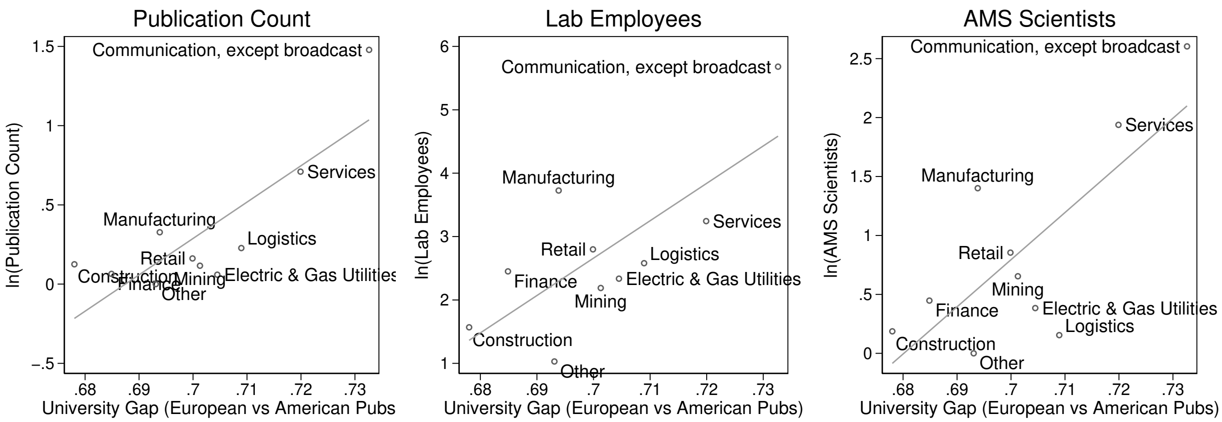


Notes: This figure compares the two scientific gap measures at the firm level. Higher values represent a larger gap between Europe and the United States. The journal citation-based gap measure (on the vertical axis) is positively correlated with the publication volume-based gap measure (on the horizontal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

Appendix Figure C3 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.

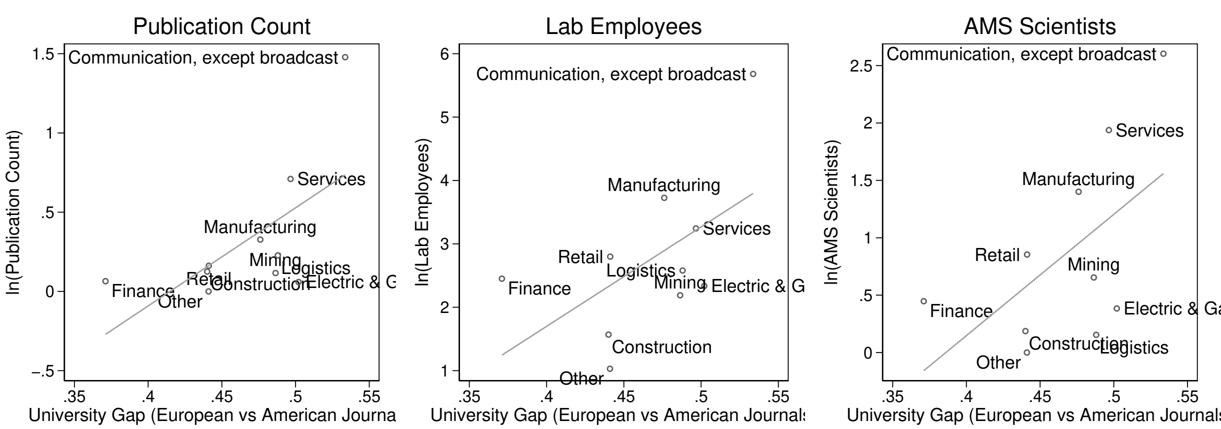
[65] It is unclear whether this represents a measurement error. The only electrical engineering journal in print during this period (1900-20) is American ("Proceedings of the Institute of Radio Engineers") and the only other electrical engineering journal indexed in the SCI before 1940 is the BELL SYSTEM 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 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 C3: 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 C4: 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 D Auxiliary Results

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 2 and 4), whereas no such relationship exists for the 1900-1919 period (Columns 1 and 3).

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

Dependent Variable	Top 10% KPST		Top 5% Cites	
	(1)	(2)	(3)	(4)
	1900-1919	1920-1940	1900-1919	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. The 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 et al. (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.

Table D2 regresses the share of university papers against measures of patent quality. We find that patents with greater technological impact (as measured through textual similarities) and forward citations are farther from university papers. Column 3 shows a standard deviation greater patent importance (Kelly et al., 2021) is associated with 1.1% fewer university papers in the 10 closest papers to patents (relative to the sample mean). This negative correlation with patent importance is around 6.4 times stronger compared to the 1000 closest paper set (Column 1). Similar contrasts are observed in Columns 2 and 4, where the negative correlation between forward patent citations and university shares is 40% larger for 10 closest papers versus 1000 closest papers to patents. The positive correlation between proxies of patent quality and corporate paper shares is consistent with corporate papers serving as inputs to technologically advanced inventions. The positive correlation between proxies of patent quality and corporate paper shares is consistent with corporate papers serving as inputs to technologically advanced inventions.

Table D2: PATENT PROXIMITY TO UNIVERSITY SCIENCE (OLS)

Dependent Variable	Univ Share (1000 Closest)		Univ Share (10 Closest)	
	(1)	(2)	(3)	(4)
Patent Importance	-0.006		-0.042	
	(0.001)		(0.002)	
Forward Patent Cites		-0.003		-0.005
		(0.000)		(0.001)
Avg of Dep Var	0.877	0.877	0.774	0.774
Year Dummies	Yes	Yes	Yes	Yes
Patent Class Dummies	Yes	Yes	Yes	Yes
R^2	0.142	0.142	0.114	0.113
Observations	658,570	658,742	658,570	658,742

Notes: The unit of analysis is a patent. The sample is limited to U.S. patents between 1926 and 1940. The dependent variable for Columns 1 and 2 is the share of university papers out of the 1000 closest papers to the focal patent; for Columns 3 and 4, it is university paper share for the 10 closest papers to the focal patent. Textual similarity between patents and papers is calculated through the SPECTER algorithm (details in Arora et al. (2023)). *Patent Importance* is calculated by dividing 10-year forward textual similarity by 5-year backward textual similarity of patents from Kelly et al. (2021). *Forward Patent Cites* refers to log of one plus 10 year forward citations received by the focal patent. All specifications include patent grant year and NBER Patent Category fixed effects.

D.1 Replication of Results with Alternative Measures of Corporate Science

Table D3: INDUSTRIAL LAB EMPLOYEES AND TECHNOLOGICAL & MARKET LEADERSHIP (POISSON)

Dependent Variable	Lab Employees							
	Technological Leadership				Market Leadership			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for Patent Cite to Science	1.872			0.622				
	(0.338)			(0.300)				
Dummy for First Patent in CPC		1.612		0.483				
		(0.196)		(0.149)				
Patent Importance			1.307	0.662				
			(0.596)	(0.559)				
ln(Assets, 1926-1930)					0.746			0.748
					(0.116)			(0.114)
Market Share						0.217		-0.444
						(0.858)		(0.796)
Dummy for Competitive Market							-0.447	-0.491
							(0.224)	(0.258)
Distance to Universities	-0.021	-0.021	0.010	0.013	-0.020	-0.017	-0.032	-0.040
	(0.039)	(0.031)	(0.034)	(0.037)	(0.021)	(0.027)	(0.032)	(0.030)
PhD Graduates	0.096	0.089	0.918	0.682	0.091	0.062	0.113	0.147
	(0.019)	(0.020)	(0.078)	(0.089)	(0.025)	(0.021)	(0.043)	(0.053)
ln(Assets)						0.750	0.738	
						(0.127)	(0.069)	
Average of Dep Var	46.012	46.012	73.721	73.721	56.157	78.168	58.817	65.549
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Pseudo-R^2	0.588	0.578	0.669	0.685	0.670	0.700	0.556	0.542
Number of Firms	465	465	423	423	355	260	389	272
Number of Obs	2,250	2,250	1,308	1,308	1,720	934	1,386	1,360

Notes: The table replicates results from Table 4 using lab size as the dependent variable. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D4: PROMINENT CORPORATE SCIENTISTS AND TECHNOLOGICAL & MARKET LEADERSHIP (POISSON)

Dependent Variable	AMS Scientists							
	Technological Leadership				Market Leadership			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy for Patent Cite to Science	2.539			1.569				
	(0.403)			(0.291)				
Dummy for First Patent in CPC		1.784		0.624				
		(0.236)		(0.113)				
Patent Importance			2.422	0.941				
			(0.497)	(0.390)				
ln(Assets, 1926-1930)					0.940			0.740
					(0.131)			(0.112)
Market Share						1.432		0.112
						(1.024)		(0.887)
Dummy for Competitive Market							-0.396	-0.316
							(0.408)	(0.402)
Distance to Universities	-0.074	-0.110	-0.107	-0.063	-0.118	-0.134	-0.135	-0.129
	(0.029)	(0.042)	(0.056)	(0.034)	(0.051)	(0.070)	(0.062)	(0.054)
PhD Graduates	0.056	0.059	0.819	0.447	0.063	0.053	0.102	0.116
	(0.011)	(0.015)	(0.111)	(0.079)	(0.014)	(0.017)	(0.046)	(0.049)
ln(Assets)						0.845	0.781	
						(0.181)	(0.086)	
Average of Dep Var	3.327	3.327	5.297	5.297	3.886	5.856	4.075	4.432
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Pseudo-R^2	0.609	0.518	0.633	0.703	0.655	0.671	0.455	0.459
Number of Firms	465	465	445	445	355	273	411	272
Number of Obs	5,564	5,564	3,116	3,116	4,277	2,378	3,663	3,536

Notes: The table replicates results from Table 4 using employment of prominent scientists as the dependent variable. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D5: CORPORATE INVESTMENT IN SCIENCE AND TYPE OF PATENT CITATIONS TO SCIENCE (POISSON)

Dependent Variable	Publication Count		Lab Employees		AMS Scientsts	
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for No Pat Cite to Sci (Base)	0.000		0.000		0.000	
	(.)		(.)		(.)	
Dummy for Pat Cite to Old Sci	-0.173		0.019		0.186	
	(0.513)		(0.457)		(0.327)	
Dummy for Pat Cite to Recent Sci	2.321		0.968		1.857	
	(0.442)		(0.298)		(0.293)	
Dummy for No Pat Cite to Sci (Base)		0.000		0.000		0.000
		(.)		(.)		(.)
Dummy for Pat Cite to Low Quality Sci		1.525		0.398		1.311
		(0.620)		(0.272)		(0.376)
Dummy for Pat Cite to High Quality Sci		2.900		1.245		1.988
		(0.371)		(0.315)		(0.310)
ln(Assets)	0.605	0.454	0.644	0.585	0.658	0.590
	(0.198)	(0.116)	(0.070)	(0.065)	(0.074)	(0.062)
Distance to Universities	0.017	0.034	-0.008	-0.004	-0.080	-0.080
	(0.030)	(0.029)	(0.029)	(0.029)	(0.032)	(0.039)
PhD Graduates	0.013	0.004	0.052	0.050	0.023	0.020
	(0.018)	(0.015)	(0.016)	(0.016)	(0.010)	(0.010)
Average of Dep Var	0.688	0.688	61.267	61.267	4.438	4.438
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.678	0.700	0.709	0.718	0.752	0.751
Number of Firms	422	422	394	394	415	415
Number of Obs	3,855	3,855	1,331	1,331	3,363	3,363

Notes: The analysis is at the firm-year level. *Dummy for Pat Cite to Recent Sci* is equal to one if the firm's patent cites a recent scientific article, and zero otherwise. A cited scientific article is "recent" if its age (grant year of the citing patent minus the publication year of the cited article) is below the average age of articles cited by patents in the same CPC and grant year. *Dummy for Pat Cite to Old Sci* is equal to one for firms whose patents cite a scientific article, but that are not recent. *Dummy for Pat Cite to High Quality Sci* is equal to one if firm patents cite scientific articles that receive above-average forward paper citations (the average is calculated for papers cited by firm patents). Standard errors are clustered at the firm level.

Table D6: Robustness Tests for Investment in Science and Technological Leadership

	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	(13) AMS Scientists
Dummy for Patent Cite to Science	0.354	0.782	1.081	1.043	0.127	1.102	0.831	-0.031	1.450	1.300	0.907	0.680	1.200
	(0.181)	(0.358)	(0.414)	(0.555)	(0.358)	(0.391)	(0.775)	(0.326)	(0.514)	(0.179)	(0.426)	(0.328)	(0.369)
Dummy for First Patent in CPC	0.258	0.068	0.093	0.254	-0.181	0.172	0.413	0.043	0.314	0.641	0.163	0.124	0.378
	(0.104)	(0.137)	(0.119)	(0.251)	(0.400)	(0.170)	(0.161)	(0.203)	(0.132)	(0.142)	(0.346)	(0.260)	(0.279)
Patent Importance	1.137	0.605	0.378	3.934	0.891	2.002	2.609	0.623	0.892	2.310	6.762	1.886	2.764
	(0.347)	(0.394)	(0.379)	(1.036)	(1.430)	(0.539)	(1.056)	(0.515)	(0.459)	(0.460)	(1.708)	(0.982)	(0.991)
ln(Assets)	0.274	0.187	0.294	0.677	0.433	0.460	0.797	0.374	0.485	0.262	0.710	0.557	0.675
	(0.059)	(0.059)	(0.068)	(0.084)	(0.101)	(0.072)	(0.177)	(0.134)	(0.150)	(0.066)	(0.126)	(0.102)	(0.074)
Distance to Universities	0.010	0.010	-0.001	0.001	0.019	-0.168	0.056	0.052	-0.195	-0.061	-0.034	-0.012	-0.097
	(0.020)	(0.025)	(0.036)	(0.105)	(0.054)	(0.125)	(0.108)	(0.036)	(0.236)	(0.054)	(0.050)	(0.032)	(0.044)
PhD Graduates	0.072	0.050	0.012	0.420	0.730	0.325	0.407	0.651	0.246	0.225	0.028	0.082	0.029
	(0.039)	(0.026)	(0.018)	(0.228)	(0.233)	(0.101)	(0.206)	(0.156)	(0.093)	(0.069)	(0.024)	(0.023)	(0.019)
Dummy for Business Group				-0.649	-0.007	0.207							
				(0.169)	(0.365)	(0.190)							
Ultimate Owner Sales HHI				-1.085	0.357	-0.072							
				(0.413)	(1.071)	(0.285)							
Average of Dep Var	0.161	0.612	0.501	2.260	295.817	11.415	3.414	252.347	19.567	0.446	0.543	50.168	3.474
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.274	0.187	0.227	0.745	0.758	0.788	0.738	0.740	0.785	0.544	0.719	0.780	0.760
Number of Firms	397	338	374	155	99	145	85	71	81	388	389	401	386
Number of Obs	2,640	907	2,246	388	93	337	587	222	538	2,536	389	401	386

Notes: The analysis is at the firm-year level for Columns 1-10 and at the firm level for Columns 11-13. 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. Details on constructing Business Group and Ultimate Owner Sales HHI are in Appendix B.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. For Columns 11-13, the continuous variables are averaged over the entire sample period (1926-1940) while the dummies take maximum value for the same period. Standard errors are clustered at the firm level. See Appendix Table B1 for details on variable construction.

Table D7: CORPORATE SCIENCE AND LEADERSHIP: SATURATED SPECIFICATION (POISSON)

Dependent Variable	Publication Count	Lab Employees	AMS Scientists
	(1)	(2)	(3)
Dummy for Patent Cite to Science	2.335	0.504	1.061
	(0.474)	(0.342)	(0.266)
Dummy for First Patent in CPC	0.843	0.671	0.440
	(0.285)	(0.206)	(0.139)
Patent Importance	-0.683	0.449	2.647
	(1.218)	(0.598)	(0.587)
ln(Assets, 1926-1930)	0.284	0.461	0.339
	(0.210)	(0.079)	(0.084)
Market Share	2.639	-0.951	0.165
	(1.170)	(0.548)	(0.655)
Dummy for Competitive Market	-1.294	-0.646	-0.486
	(0.601)	(0.233)	(0.385)
Distance to Universities	0.066	0.005	-0.086
	(0.043)	(0.033)	(0.040)
PhD Graduates	0.299	0.530	0.449
	(0.183)	(0.089)	(0.112)
Average of Dep Var	1.137	105.332	7.092
Year Dummies	Yes	Yes	Yes
3-Digit SIC Dummies	No	No	No
Pseudo-R^2	0.677	0.702	0.739
Number of Firms	272	241	264
Number of Obs	2,371	798	2,021

Notes: The analysis is at the firm-year level. "ln(Assets, 1926-1930)" takes the natural log average assets for 1926-1930. Standard errors are clustered at the firm level. See Appendix Table B1 for details on variable construction.

Table D7 estimates a conditional Poisson specification which includes all level variables from Tables 4, D3, and D4.

Table D8: INDUSTRIAL RESEARCH LABS AND GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	Lab Employees							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap	15.139	26.018	11.034	8.385	13.312	-19.793	12.065	26.634
	(4.600)	(6.230)	(5.180)	(5.133)	(5.725)	(73.854)	(6.373)	(4.223)
Univ Gap × Distance to Universities		-1.629						
		(0.602)						
Univ Gap × Dummy for Patent Cite to Science			14.387					
			(12.150)					
Univ Gap × Dummy for First Patent in CPC				15.395				
				(7.971)				
Univ Gap × Patent Importance					13.607			
					(24.728)			
Univ Gap × ln(Assets, 1926-1930)						1.569		
						(3.604)		
Univ Gap × Market Share							56.913	
							(32.873)	
Univ Gap × Dummy for Competitive Market								-21.615
								(5.592)
Dummy for Patent Cite to Science			-9.252					
			(8.426)					
Dummy for First Patent in CPC				-10.050				
				(5.537)				
Patent Importance					-8.427			
					(17.043)			
ln(Assets, 1926-1930)						-0.412		
						(2.524)		
Market Share							-40.320	
							(23.029)	
Dummy for Competitive Market								14.566
								(3.932)
ln(Assets)	0.713	0.719	0.633	0.608	0.504		0.743	0.709
	(0.071)	(0.072)	(0.069)	(0.069)	(0.063)		(0.107)	(0.053)
Distance to Universities	-0.025	1.124	-0.006	-0.009	0.012	-0.024	-0.026	-0.045
	(0.029)	(0.424)	(0.028)	(0.025)	(0.025)	(0.024)	(0.029)	(0.038)
PhD Graduates	0.073	0.069	0.055	0.050	0.462	0.097	0.065	0.110
	(0.020)	(0.019)	(0.017)	(0.016)	(0.090)	(0.026)	(0.021)	(0.030)
Average of Dependent Variable	61.267	61.267	61.267	61.267	82.962	56.157	78.168	58.817
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.698	0.700	0.717	0.718	0.738	0.681	0.718	0.610
Number of Firms	394	394	394	394	338	355	260	389
Number of Observations	1,331	1,331	1,331	1,331	934	1,720	934	1,386

Notes: The table replicates results from Table 5 using lab size as the dependent variable. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D9: PROMINENT CORPORATE SCIENTISTS AND GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	AMS Scientists							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap	16.774	26.934	7.220	4.188	12.592	-112.929	3.348	27.651
	(5.759)	(9.878)	(4.445)	(5.237)	(6.388)	(96.160)	(6.941)	(4.995)
Univ Gap × Distance to Universities		-1.656						
		(1.139)						
Univ Gap × Dummy for Patent Cite to Science			39.396					
			(9.088)					
Univ Gap × Dummy for First Patent in CPC				30.714				
				(7.661)				
Univ Gap × Patent Importance					64.562			
					(22.304)			
Univ Gap × ln(Assets, 1926-1930)						6.227		
						(4.594)		
Univ Gap × Market Share							102.786	
							(30.930)	
Univ Gap × Dummy for Competitive Market								-24.537
								(7.597)
Dummy for Patent Cite to Science			-25.680					
			(6.381)					
Dummy for First Patent in CPC				-20.321				
				(5.395)				
Patent Importance					-43.109			
					(15.424)			
ln(Assets, 1926-1930)						-3.546		
						(3.276)		
Market Share							-70.589	
							(21.938)	
Dummy for Competitive Market								16.393
								(5.199)
ln(Assets)	0.851	0.853	0.629	0.706	0.601		0.765	0.760
	(0.110)	(0.110)	(0.066)	(0.092)	(0.071)		(0.112)	(0.086)
Distance to Universities	-0.134	1.041	-0.068	-0.093	-0.120	-0.118	-0.160	-0.165
	(0.060)	(0.821)	(0.030)	(0.050)	(0.066)	(0.046)	(0.077)	(0.075)
PhD Graduates	0.058	0.057	0.023	0.024	0.394	0.068	0.050	0.100
	(0.014)	(0.014)	(0.010)	(0.010)	(0.107)	(0.014)	(0.015)	(0.026)
Average of Dependent Variable	4.438	4.438	4.438	4.438	6.063	3.886	5.856	4.075
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.687	0.688	0.764	0.727	0.758	0.686	0.705	0.593
Number of Firms	415	415	415	415	374	355	273	411
Number of Observations	3,363	3,363	3,363	3,363	2,263	4,277	2,378	3,663

Notes: The table replicates results from Table 5 using employment of prominent scientists as the dependent variable. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D10: 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$_{t-1}$)	0.313			0.289	0.187			0.227
	(0.118)			(0.276)	(0.100)			(0.093)
ln(Lab Employees$_{t-1}$)		-0.142		-0.237		0.075		0.020
		(0.064)		(0.061)		(0.049)		(0.045)
ln(AMS Scientists$_{t-1}$)			0.218	0.247			0.204	0.072
			(0.160)	(0.171)			(0.093)	(0.068)
ln(Patent Stock$_{t-1}$)	0.165	0.447	0.169	0.188	0.703	0.912	0.673	0.755
	(0.095)	(0.138)	(0.116)	(0.142)	(0.086)	(0.083)	(0.083)	(0.091)
Distance to Universities	-0.092	-0.077	-0.070	-0.055	-0.026	-0.009	0.001	-0.014
	(0.067)	(0.079)	(0.056)	(0.064)	(0.026)	(0.019)	(0.022)	(0.024)
PhD Graduates$_{t-1}$	0.092	0.042	0.089	0.070	0.065	0.030	0.064	0.037
	(0.040)	(0.044)	(0.040)	(0.054)	(0.015)	(0.018)	(0.015)	(0.018)
ln(Assets$_{t-1}$)	0.927	1.146	0.964	1.105	-0.103	-0.126	-0.090	-0.164
	(0.185)	(0.187)	(0.205)	(0.225)	(0.087)	(0.096)	(0.098)	(0.100)
Average of Dep Var	1.574	1.563	1.499	1.563	0.844	0.799	0.802	0.799
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.798	0.781	0.790	0.799	0.712	0.738	0.702	0.746
Number of Firms	321	308	317	308	418	384	415	384
Number of Obs	1,888	497	1,754	497	3,916	960	3,589	960

Notes: The analysis is at the firm-year level. $ln(Assets_{t-1})$, $ln(PatentStock_{t-1})$, $ln(PublicationStock_{t-1})$ and $ln(LabEmployees_{t-1})$ take the natural log of lagged assets, patent stock, publication stock and lab employees respectively. Patent and publication stock are calculated using a perpetual inventory method with a 15% rate of depreciation. The dependent variable for Columns 1-4 is the number of firm patents in the top 5% of stock market value (Kogan et al., 2017). The dependent variable for Columns 5-8 is the number of firm patents in the top 5% in terms of forward citations. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

D.2 Replication of Results with Journal Citation-Based Measure

Table D11: CORPORATE PUBLICATIONS AND JOURNAL CITATION-BASED GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	Publication Count							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap (Cites)	4.960	7.334	-1.080	-0.013	1.953	-39.378	-1.277	9.788
	(2.855)	(4.579)	(1.960)	(2.478)	(3.049)	(32.355)	(2.361)	(2.031)
Univ Gap (Cites) × Distance to Universities		-0.310						
		(0.326)						
Univ Gap (Cites) × Dummy for Patent Cite to Science			31.922					
			(7.034)					
Univ Gap (Cites) × Dummy for First Patent in CPC				12.626				
				(4.029)				
Univ Gap (Cites) × Patent Importance					46.339			
					(14.924)			
Univ Gap (Cites) × ln(Assets, 1926-1930)						2.180		
						(1.628)		
Univ Gap (Cites) × Market Share							59.715	
							(28.968)	
Univ Gap (Cites) × Dummy for Competitive Market								-10.522
								(4.347)
Dummy for Patent Cite to Science			-13.659					
			(3.459)					
Dummy for First Patent in CPC				-4.611				
				(1.967)				
Patent Importance					-18.339			
					(6.919)			
ln(Assets, 1926-1930)						-0.112		
						(0.853)		
Market Share							-24.045	
							(14.643)	
Dummy for Competitive Market								4.105
								(2.014)
ln(Assets)	0.839	0.843	0.741	0.678	0.678		0.580	0.806
	(0.215)	(0.215)	(0.105)	(0.142)	(0.107)		(0.069)	(0.098)
Distance to Universities	-0.011	0.155	0.041	0.028	-0.020	-0.042	-0.027	-0.073
	(0.031)	(0.178)	(0.028)	(0.025)	(0.059)	(0.043)	(0.053)	(0.057)
PhD Graduates	0.080	0.078	0.010	0.022	0.490	0.071	0.040	0.140
	(0.032)	(0.034)	(0.015)	(0.020)	(0.101)	(0.026)	(0.021)	(0.092)
Average of Dependent Variable	0.688	0.688	0.688	0.688	0.939	0.628	0.922	0.624
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.640	0.640	0.725	0.684	0.737	0.646	0.746	0.404
Number of Firms	422	422	422	422	397	355	273	417
Number of Observations	3,855	3,855	3,855	3,855	2,640	4,860	2,747	4,252

Notes: The table replicates results from Table 5 using the journal citation-based measure of university gap. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D12: INDUSTRIAL RESEARCH LABS AND JOURNAL CITATION-BASED GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	Lab Employees							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap (Cites)	2.514	5.620	0.039	0.954	2.574	-40.384	2.085	5.785
	(1.580)	(2.655)	(2.578)	(1.808)	(2.068)	(21.022)	(2.187)	(2.016)
Univ Gap (Cites) × Distance to Universities		-0.441						
		(0.308)						
Univ Gap (Cites) × Dummy for Patent Cite to Science			12.687					
			(5.040)					
Univ Gap (Cites) × Dummy for First Patent in CPC				4.139				
				(2.117)				
Univ Gap (Cites) × Patent Importance					7.739			
					(6.924)			
Univ Gap (Cites) × ln(Assets, 1926-1930)						2.098		
						(1.065)		
Univ Gap (Cites) × Market Share							12.472	
							(18.868)	
Univ Gap (Cites) × Dummy for Competitive Market								-9.111
								(2.975)
Dummy for Patent Cite to Science			-5.064					
			(2.213)					
Dummy for First Patent in CPC				-1.185				
				(0.940)				
Patent Importance					-2.289			
					(2.844)			
ln(Assets, 1926-1930)						-0.304		
						(0.572)		
Market Share							-5.977	
							(9.095)	
Dummy for Competitive Market								3.647
								(1.373)
ln(Assets)	0.731	0.738	0.651	0.629	0.515		0.754	0.766
	(0.075)	(0.075)	(0.060)	(0.064)	(0.062)		(0.110)	(0.071)
Distance to Universities	-0.018	0.200	0.003	-0.006	0.015	-0.024	-0.018	-0.032
	(0.025)	(0.152)	(0.026)	(0.024)	(0.023)	(0.025)	(0.027)	(0.036)
PhD Graduates	0.072	0.068	0.048	0.051	0.466	0.100	0.065	0.109
	(0.021)	(0.020)	(0.015)	(0.016)	(0.086)	(0.029)	(0.022)	(0.033)
Average of Dependent Variable	61.267	61.267	61.267	61.267	82.962	56.157	78.168	58.817
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.692	0.693	0.727	0.713	0.737	0.682	0.709	0.596
Number of Firms	394	394	394	394	338	355	260	389
Number of Observations	1,331	1,331	1,331	1,331	934	1,720	934	1,386

Notes: The table replicates results from Table D8 using the journal citation-based measure of university gap. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D13: PROMINENT CORPORATE SCIENTISTS AND JOURNAL CITATION-BASED GAP IN UNIVERSITY SCIENCE (POISSON)

Dependent Variable	AMS Scientists							
	Baseline	Univ Distance	Technological Leadership			Market Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Univ Gap (Cites)	3.465	4.875	-0.820	1.108	2.708	-54.337	-0.489	5.592
	(1.999)	(3.583)	(1.517)	(1.678)	(1.851)	(27.319)	(2.649)	(2.362)
Univ Gap (Cites) × Distance to Universities		-0.211						
		(0.404)						
Univ Gap (Cites) × Dummy for Patent Cite to Science			21.790					
			(3.168)					
Univ Gap (Cites) × Dummy for First Patent in CPC				7.356				
				(1.818)				
Univ Gap (Cites) × Patent Importance					25.456			
					(6.835)			
Univ Gap (Cites) × ln(Assets, 1926-1930)						2.794		
						(1.331)		
Univ Gap (Cites) × Market Share							42.878	
							(20.475)	
Univ Gap (Cites) × Dummy for Competitive Market								-8.982
								(4.079)
Dummy for Patent Cite to Science			-8.317					
			(1.500)					
Dummy for First Patent in CPC				-2.289				
				(0.859)				
Patent Importance					-9.661			
					(3.158)			
ln(Assets, 1926-1930)						-0.465		
						(0.676)		
Market Share							-19.435	
							(10.071)	
Dummy for Competitive Market								3.569
								(1.807)
ln(Assets)	0.869	0.870	0.687	0.725	0.626		0.823	0.832
	(0.117)	(0.116)	(0.064)	(0.089)	(0.073)		(0.141)	(0.104)
Distance to Universities	-0.121	-0.014	-0.055	-0.092	-0.121	-0.112	-0.126	-0.147
	(0.057)	(0.238)	(0.027)	(0.051)	(0.067)	(0.047)	(0.065)	(0.070)
PhD Graduates	0.061	0.061	0.015	0.028	0.419	0.072	0.056	0.100
	(0.014)	(0.015)	(0.009)	(0.010)	(0.106)	(0.014)	(0.017)	(0.028)
Average of Dependent Variable	4.438	4.438	4.438	4.438	6.063	3.886	5.856	4.075
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Pseudo-R^2	0.677	0.678	0.776	0.717	0.756	0.680	0.691	0.580
Number of Firms	415	415	415	415	374	355	273	411
Number of Observations	3,363	3,363	3,363	3,363	2,263	4,277	2,378	3,663

Notes: The table replicates results from Table D9 using the journal citation-based measure of university gap. The analysis is at the firm-year level. See Appendix Table B1 for details on variable construction. Standard errors are clustered at the firm level.

Table D14: CORPORATE SCIENCE AND STOCK MARKET VALUE (JOURNAL CITATION-BASED GAP) (OLS)

Dependent Variable	ln(Market Capitalization)			
	University Gap Split		University Gap Split for Subsample with Lab & AMS Scientist	
	(1) University Gap Below Mean	(2) University Gap Above Mean	(3) University Gap Below Mean	(4) University Gap Above Mean
ln(Publication Stock$_{t-1}$)	0.239	0.212	0.210	0.247
	(0.090)	(0.072)	(0.118)	(0.079)
ln(Patent Stock$_{t-1}$)	0.211	0.016	0.110	-0.197
	(0.043)	(0.043)	(0.084)	(0.064)
ln(Assets$_{t-1}$)	0.724	0.794	0.767	0.965
	(0.074)	(0.068)	(0.151)	(0.080)
Average of Dependent Variable	19.567	19.248	20.042	19.984
Year Fixed Effects	Yes	Yes	Yes	Yes
3-Digit SIC Dummies	Yes	Yes	Yes	Yes
R^2	0.727	0.798	0.746	0.906
Number of Firms	167	154	85	54
Number of Observations	1,776	1,557	173	96

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 D15: CORPORATE SCIENCE AND MARKET-TO-BOOK RATIOS (OLS)

Dependent Variable	ln(Tobin's Q)				
		Publication-Based Gap Split		Journal Citation-Based Gap Split	
	(1) All	(2) University Gap Below Mean	(3) University Gap Above Mean	(4) University Gap Below Mean	(5) University 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 are included in all columns. Standard errors are clustered at the firm level.

References for Appendices

Arora, A., Belenzon, S., Cioaca, L., Sheer, L., & Zhang, H. (2023). The Effect of Public Science on Corporate R&D. *Working Paper*.

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., & Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, *111*(3), 871–98.

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. *Lessons from the Quota Acts*.

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.