FRONTIER KNOWLEDGE AND SCIENTIFIC PRODUCTION: EVIDENCE FROM THE COLLAPSE OF INTERNATIONAL SCIENCE*

Alessandro Iaria
Carlo Schwarz
Fabian Waldinger

We show that World War I and the subsequent boycott against Central scientists severely interrupted international scientific cooperation. After 1914, citations to recent research from abroad decreased and paper titles became less similar (evaluated by latent semantic analysis), suggesting a reduction in international knowledge flows. Reduced international scientific cooperation led to a decline in the production of basic science and its application in new technology. Specifically, we compare productivity changes for scientists who relied on frontier research from abroad, to changes for scientists who relied on frontier research from home. After 1914, scientists who relied on frontier research from abroad published fewer papers in top scientific journals, produced less Nobel Prize–nominated research, introduced fewer novel scientific words, and introduced fewer novel words that appeared in the text of subsequent patent grants. The productivity of scientists who relied on top 1% research declined twice as much as the productivity of scientists who relied on top 3% research. Furthermore, highly prolific scientists experienced the starkest absolute productivity declines. This suggests that access to the very best research is key for scientific and technological progress. *JEL Codes:* O3, N3, N4, O31, O5, N30, N40, J44, I23.

*We thank the editor, Andrei Shleifer, and four anonymous referees for very useful comments and suggestions. We thank M. Z. Bhatti, A. Buensuceso, G. Daubaris, T. Decker, J. Dieminger, T. Dimitrov, L. Essa, F. Filsinger, D. Full, F. Ganz, C. Kuehn, A.-K. Marko, L. Mergele, T. Nowacki, J. Ostendorf, N. Porschke, J. Prestwich, A. Reitz, Constanze Schwarz, C. Vorrath, and S. Wies for excellent research assistance. We thank Ran Abramitzky, David Card, Raj Chetty, James Fenske, Claudia Goldin, Larry Katz, Fabio Montobbio, Petra Moser, Roland Rathelot, Paul Rhode, Bruce Weinberg, and seminar participants at Berkeley, Bonn, Columbia, Copenhagen, CREST, EBRD, EEA, EPFL, Essex, EUI, Georgia Tech, HBS, LEI-Brick Turin, LSE, MIT, Munich, NBER SI, NYU, Ohio State, Oxford, PSE, Queens University, Stanford, UCLA, Warwick, and Zurich for excellent comments and suggestions. Waldinger gratefully acknowledges start-up funding from the Sloan Research Project on the Economics of Knowledge Contribution and Distribution and funding from ERC Starting Grant no. 335573.

The Quarterly Journal of Economics (2018), 927–991. doi:10.1093/qje/qjx046.
Advance Access publication on January 15, 2018.

I. Introduction

The creation of ideas is crucial for scientific progress, technological innovation, and economic development, particularly in a world where "knowledge has taken over much of the economy" (The Economist 2000). As argued by many scholars (e.g., Arrow 1962; Mokyr 2002), one of the major inputs in the creation of new ideas is existing knowledge. Most famously, Isaac Newton acknowledged the importance of existing knowledge in his letter to Robert Hooke: "If I have seen further, it is by standing on ye shoulders of Giants" (Newton 1675). The quote not only emphasizes that scientists build on existing knowledge to produce new ideas, but also that knowledge produced by scientific "giants," that is, frontier knowledge, is particularly important. Access to existing knowledge not only fuels basic scientific progress but is also key for the development of new technologies, as emphasized by theoretical models of economic growth (e.g., Romer 1986, 1990; Jones 1995; Weitzman 1998).

In the first part of the article, we document a sharp decline in international scientific cooperation around World War I (WWI). This decline severely reduced international citations in scientific papers, including citations to the international knowledge frontier. In the second part of the article, we study how reduced access to the international knowledge frontier affected the production of basic science and its application in new technologies.

With the beginning of the war, the world split into the Allied (United Kingdom, France, later the United States, and a number of smaller countries) and Central (Germany, Austria-Hungary, Ottoman Empire, Bulgaria) camps. The involvement of scientists in the war effort and the extremely nationalistic stance taken by many scientists in support of their homeland, Germany in particular, pitted scientists in the two camps against each other. We document that the delivery of international journals was severely delayed and that international conferences were canceled or only involved scientists from one of the warring camps. Allied scientists were cut off from their peers in Central countries; in particular from Germany, a country whose scientists had received more than 40% of Nobel Prizes in physics and chemistry in the prewar period. Similarly, Central scientists were cut off from their peers in Allied countries; in particular from the United Kingdom (20% of Nobel Prizes), France (15% of Nobel Prizes), and the United States, the rising scientific superpower. This schism of the scientific world persisted during the postwar years because Allied scientists organized a boycott against Central scientists to punish them for their involvement in the war effort.

To quantify the decline in international scientific cooperation and to measure how it affected scientific progress, we collect data from various historical sources. First, we digitize more than 60,000 individual records from *Minerva—Handbuch der Gelehrten Welt*, the most comprehensive worldwide listing of university professors for this period, and we compile two censuses of all university scientists in the world for 1900 and 1914. Second, we collect data on all scientific publications, including references, in 160 top scientific journals for the period 1900 to 1930 from the *ISI Web of Science*. Third, we collect data on all Nobel Prize nominations for the years 1905 to 1945 from the Nobel archives. Fourth, we collect data on more than 2.5 million U.S. patents.

In the first part of the article, we show that international citations in scientific papers severely declined during WWI and the subsequent boycott against Central scientists. After 1914, papers contained fewer citations to recent research from outside the camp, relative to research from home, that is, Allied papers contained fewer citations to Central research, and Central papers contained fewer citations to Allied research. We estimate that the share of citations to research from outside the camp fell by 0.22, a decline of about 85%. We find a smaller decline in relative citations to foreign research from inside the camp, consistent with a smaller interruption of international scientific cooperation.

Moreover, we explore whether WWI and the boycott also affected citations to top research by focusing on references quoting research that ended up in the top percentiles of the citation distribution. After 1914, citations to top 5%, top 3%, and even top 1% research from outside the camp declined, relative to citations to the corresponding top research from home.

The observed changes in international citations could be caused either by scientists not knowing about recent foreign research or by scientists deciding not to cite foreign research for political reasons. To distinguish between these two possibilities, we explore citations to prewar research. In contrast to recent research, international citations to prewar research did not fall disproportionately after 1914. This suggests that the observed changes in international citations were presumably caused by scientists not knowing about recent foreign research.

In further results, we analyze how the breakdown in international scientific cooperation affected the similarity of papers produced in the different camps. We use the machine learning algorithm latent semantic analysis (Deerwester et al. 1990; Landauer, Foltz, and Laham 1998) to measure the similarity of paper titles. After 1914, the similarity to papers from outside the camp fell by 0.5 standard deviations relative to the similarity to papers from home. The similarity to papers from inside the camp did not fall significantly. These results suggest that the breakdown in international scientific cooperation also led to a divergence of research in the two camps.

In the second part of the article, we study consequences of the decline in international scientific cooperation for the production of basic science and its application in new technologies. Specifically, we compare yearly productivity changes of scientists in field-country pairs that, in the prewar period, relied on frontier research from abroad, for example, biochemists in the United States, to scientists in field-country pairs that relied on frontier research from home, for example, biologists in the United States. After 1914, scientists who relied on frontier research (as measured by the top 1%) from outside the camp, rather than from home, published significantly fewer papers. The results imply that U.S. biochemists published 0.1 standard deviations fewer papers per year after 1914, a productivity reduction of about 30%, compared to U.S. biologists. We also show that productivity declined for scientists in field-country pairs that relied on frontier research from inside the camp, but not significantly so.

Further results indicate that scientists who relied on top 1% research experienced productivity reductions that were at least twice as large as those of scientists who relied on top 3% or top 5% research. While researchers have always grasped the relevance of frontier research, our results emphasize the narrow-edged nature of the knowledge frontier.

We investigate whether the relative changes in productivity were most likely caused by a reduction in international knowledge flows, or by more general disruption during WWI. To control for disruption that affected all scientists to the same extent, all regressions include year fixed effects. Results remain unchanged if we control for camp-times-year, field-times-year, or camp-times-field-times-year fixed effects. These additional fixed effects control for war-related and other changes that differentially affected scientists in different camps (e.g., Allied scientists), fields (e.g., chemists), or fields within camps (e.g., Allied chemists). We also estimate regressions that include various measures of war intensity, such as the number of total or civilian deaths. To further probe whether WWI differentially affected death rates of scientists, we collect data on more than 6,500 obituaries published in contemporary scientific journals. In general, scientists in our sample did not die disproportionately during WWI. Moreover, we show that scientists reliant on frontier research from abroad did not die disproportionately during this period. Additionally, we show that the results are robust to excluding chemists, who were most heavily involved in weapons development, and to considering only publications in home-camp journals, since publishing opportunities in foreign journals may have dwindled.

We also investigate effects on three alternative measures of scientific productivity: scientific breakthroughs, new scientific concepts, and new scientific concepts with technological applications. We find that scientists who relied heavily on frontier research from outside the camp, rather than from home, produced fewer scientific breakthroughs, as measured by research worthy of a Nobel Prize nomination.

We also study effects on new scientific concepts. Scientists who relied heavily on frontier research from outside the camp, rather than from home, produced fewer papers that introduced novel words, which serve as a measure of new scientific concepts. We define novel words as words that the scientist first used in a title of a paper published between 1905 and 1930 and that had not been used in any prior paper title. Examples of words that were introduced in this period are electroencephalogram, magnetron, hormone, isotope, and superconductor.

Furthermore, we study effects on the technological application of basic science. We develop a text-based method to establish a link from basic science to technology.[1] We search the full text of more than 2.5 million U.S. patents, containing 7.6 billion words, for the novel scientific words that scientists introduced in this

1. Many scientific advances that affect the development of new technology are not formally cited in patents. For example, U.S. patent no. 3,699,947 "Electroencephalograph Monitoring Apparatus," granted in 1972, does not mention any scientific paper, not even those of Hans Berger, who laid the scientific foundations of electroencephalography in the 1920s and 1930s. Our text-based method also allows us to measure effects on technology for a time period before the U.S. Patent Office introduced formal citations to basic science (in 1947).

period. For example, the novel scientific word "electroencephalogram" appeared seven times in subsequent patents and "magnetron" appeared 9,638 times. The measure captures connections between science and technology even if patents do not cite the relevant scientific papers. We find that scientists who relied heavily on frontier research from outside the camp, rather than from home, introduced fewer innovative words that found applications in patents.

Finally, we show that access to frontier research did not affect all scientists to the same extent. Output of above-median productivity scientists decreased 5 to 15 times more, in absolute terms, than output of below-median scientists. These results suggest a complementarity between access to frontier research and the underlying quality of scientists.

Our findings contribute to the literature on the effect of basic science on technological development, a link that is difficult to establish empirically. Our results indicate that access to frontier knowledge impacts the production of basic science that is applied in the development of new technology. Other research has shown that increased funding from the U.S. National Institutes of Health (NIH) for basic biomedical research increases patenting by private sector companies (Azoulay et al. 2016) and that NIH open access mandates increase citations to biomedical research by inventors (Bryan and Ozcan 2016).[2]

Our findings emphasize that access to existing frontier research is particularly important for the creation of ideas and that high-quality scientists make greater use of it. Because the physical costs of gaining access to frontier research have fallen since the early twentieth century, especially with the introduction of the Internet and improved transportation, the main cost of access today lies in discerning the knowledge frontier from the millions of scientific papers published every year. While not specifically investigating the role of frontier knowledge, previous literature has shown that access to existing knowledge affects follow-on research. For

2. Increased funding for universities and the establishment of technical universities increases patenting (Aghion et al. 2009; Toivanen and Väänänen 2016). Earlier research shows that basic science is associated with private sector innovation, without taking advantage of plausibly exogenous variation in basic science (e.g., Jaffe 1989; Adams 1990; Acs, Audretsch, and Feldman 1992; Mansfield 1995). More broadly, universities are associated with faster growth (Valero and Van Reenen 2016), and engineers are particularly important (Murphy, Shleifer, and Vishny 1991).

example, materials that have been deposited in biological resource centers, which collect and distribute biological material, are more likely to be used in follow-on research (Furman and Stern 2011). Intellectual property rights increase the cost of using prior knowledge in follow-on research (Scotchmer 1991; Murray et al. 2009; Williams 2013; Biasi and Moser 2015; Galasso and Schankerman 2015). The compulsory licensing of German patents after WWI, for example, increased patenting by U.S. inventors in the 1930s (Moser and Voena 2012).

Our results also contribute to the literature on the knowledge production function by highlighting the importance of frontier knowledge. The existing literature has shown that papers that cite "atypical combinations" of references are more likely to become a "hit" (Uzzi et al. 2013; Wang, Veugelers, and Stephan 2016), as are papers that predominantly cite recent as well as some older references (Mukherjee et al. 2017). More generally, human capital is more important for scientific production than physical capital (Waldinger 2016). Star scientists are key, because they affect the productivity of coauthors (Azoulay et al. 2010; Oettl 2012; Borjas and Doran 2015), attract other good scientists to their universities (Waldinger 2016, Agrawal, McHale, and Oettl 2017), attract researchers to promising research fields (Moser, Voena, and Waldinger 2014), and train PhD students (Waldinger 2010).[3] With the stock of knowledge constantly increasing, scientists must absorb ever more information to reach the knowledge frontier; therefore, they must invest more time in training and collaborate in larger teams (Wuchty, Jones, and Uzzi 2007; Jones 2009).

The results also speak to the literature on international knowledge flows by showing that political events can disrupt international knowledge flows and lower scientific productivity. Previous research has shown that city, state, and country borders are important barriers to knowledge flows, as measured by patent citations (e.g., Jaffe, Trajtenberg, and Henderson 1993; Peri 2005; Thompson and Fox-Kean 2005; Belenzon and Schankerman 2013; Head, Li, and Minondo 2015). Reductions in travel costs boost collaborations of scientists in different cities (Catalini, Fons-Rosen, and Gaule 2016). Book translations from languages

3. Other research has shown negative effects of stars when journal and faculty slots are fixed (Borjas and Doran 2012). Similarly, star scientists do not seem to have a positive effect on their peers in the same department (Waldinger 2012; Borjas and Doran 2015; Agrawal, McHale, and Oettl 2017).

spoken in Western countries (e.g., English) to languages spoken in Communist countries (e.g., Russian) were rare during the Cold War period but increased substantially after the collapse of the Soviet Union (Abramitzky and Sin 2014).

II. A Shock to International Scientific Cooperation

II.A. A Brief History of Science around WWI

Science became increasingly international during the second half of the nineteenth century, particularly in the years leading up to WWI—the so-called golden age of internationalism in science (Crawford 1988). Scientists published their most important contributions in international journals, conferences became more international, and scientific societies increased international cooperation. In 1899, leading nations founded the International Association of Academies to "facilitate scientific intercourse between the different countries" (Greenaway 1996). To improve access to international research, the Royal Society, the oldest scientific society in the world, coordinated the publication of the *International Catalogue of Scientific Literature,* which collected the titles of virtually all scientific papers and facilitated the search for these papers with English, German, French, and Italian tables of contents.

The increasing internationalization of science was abruptly interrupted by the outbreak of WWI, at the end of July 1914. The Western world split into two warring camps with the Allies (UK, France, later the United States, and a number of smaller countries) fighting the Central Powers (Germany, Austria-Hungary, the Ottoman Empire, and Bulgaria) (see Table I). While the war caused millions of military deaths, it caused relatively few civilian casualties in the major scientific powers (United States: 757 deaths, United Kingdom: 16,829, mostly merchant fleet, and Germany: 720), because the war was not fought on the territories of these countries.

All major war participants enlisted some of their most prominent scientists to support the war effort, particularly for the development of chemical weapons. The German unit was led by future Nobel Prize laureate Fritz Haber, who assembled a team of prominent chemists to develop new poisonous gases. His team included seven future Nobel laureates: James Franck, Gustav Hertz, Otto Hahn, Walter Nernst, Emil Fischer, Heinrich Wieland, and

TABLE I
SCIENTIFIC CAMPS DURING WWI AND THE BOYCOTT

Allies	Centrals	Neutrals
United States	Germany	Switzerland
United Kingdom (incl. Ireland)	Austria	Netherlands
France	Hungary	Sweden
Canada	Bulgaria	Denmark
Japan	Ottoman E./Turkey	Norway
Italy		Czechoslovakia
Belgium		Finland
Australia		Spain
Romania		Monaco
Poland		
Brazil		
South Africa		
Greece		
New Zealand		
Portugal		
Serbia		

Notes. Countries are classified following the definition of the International Research Council (IRC) and ordered by scientific output in our data. Austria-Hungary was split into two countries after WWI. Czechoslovakia was part of Austria-Hungary before WWI and became a sovereign state after 1918. Turkey emerged from parts of the Ottoman Empire after WWI.

Richard Willstätter (Van der Kloot 2004). The French unit was led by Victor Grignard, who had received the Nobel Prize in 1912. The U.S. unit also enlisted prominent scientists, including the future president of Harvard University, James Bryant Conant.

During this period, many scientists, particularly those from Germany, took a nationalistic stance and even issued statements in support of their home country's military actions. In the infamous *Manifesto of the 93*, which was widely published in October 1914, 93 German intellectuals, among them 14 science Nobel laureates, declared their support for Germany's military actions, the killing of Belgian civilians, and the destruction of Leuven with its famous university library. Two weeks later, 3,000 German university teachers endorsed a declaration that "Europe's culture depends on the victory of the German military" (Reinbothe 2006, 99). In a reply that was published in *Nature*, the British chemist and Nobel laureate William Ramsay condemned German scientists, stating that "their ideal . . . is to secure world supremacy for their race" (Ramsay 1914).

Scientists' participation in the war effort and their hostile attitude toward their international peers soured international scientific relations. As early as October 1914, Ramsay had suggested "restrictions of the Teutons" (Ramsay 1914) for the postwar era. Just before the end of the war, Allied scientists organized a conference that paved the way for a boycott against Central scientists. The scientists announced that:

> The Allied Nations are forced to declare that they will not be able to resume personal relations in scientific matters with their enemies until the Central Powers can be readmitted into the concert of civilized nations. (quoted in Lehto 1998, 18)

At a follow-up conference, more than 200 scientists from 12 Allied countries founded the International Research Council (IRC) to organize postwar international scientific cooperation.[4] The IRC ensured that scientists from Central countries were effectively cut off from Allied scientific associations and international scientific meetings, even if the associations or conference organizers were not officially affiliated with the IRC (Schroeder-Gudehus 1973). While the boycott was strictly enforced in the first postwar years, its strength declined over time. In 1922, the Allied majority rejected a proposal by Neutral scientists to invite Central scientists to join the IRC (Cock 1983; Lehto 1998, 38). In the following years, the Allied position softened, and the boycott was officially terminated in June 1926 (Lehto 1998, 40).[5] Two years later, the eminent German mathematician David Hilbert was honored to deliver the opening address of the International Congress of Mathematicians in Bologna. He proclaimed:

> It makes me very happy that after a long, hard time all the mathematicians of the world are represented here. This is as it should be and as it must be for the prosperity of our beloved science ... For mathematics, the whole cultural world is a single country. (quoted in Reid 1970, 188)

II.B. Delivery of International Journals and Attendance of Conferences

During the war and the subsequent boycott, both Allies and Centrals became increasingly strict about sharing scientific

4. The IRC replaced the International Association of Academies that had overseen international scientific relations in the prewar era. The IRC statutes explicitly excluded former Central countries, but some formerly Neutral countries were invited to join as members (Kevles 1971, 58).

5. In June 1926, Germany, Austria, Hungary, and Bulgaria were invited to join the IRC. While the German scientific academies officially declined the invitation, the boycott was effectively terminated at this point.

knowledge with foreign countries. Access to foreign journals became restricted and most international conferences were canceled during the war. Central scientists were banned from attending international conferences during the postwar boycott. More generally, most efforts to foster international scientific cooperation were interrupted during this period. The publication of the *International Catalogue of Scientific Literature*, for example, was discontinued after 1914.

1. Access to Scientific Journals from Foreign Countries. We measure how the war and the boycott reduced access to foreign journals by investigating entry stamps from the Harvard library. To register the delivery of a journal, Harvard librarians placed an entry stamp on each issue upon arrival (see Online Appendix Figure A.1 for an example). We collect data on these stamps for 1910, 1913, 1917, 1919, 1921, 1923, and 1927 for four international journals: the *Zeitschrift für Analytische Chemie*, the *Annalen der Physik*, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, and *Nature*. We then calculate the average delay between the publication of a journal and its arrival at Harvard (see Online Appendix E.1 for details).

Before the war, the German *Zeitschrift für Analytische Chemie* arrived with a delay of about 26 days (Figure I, Panel A). By 1917, the delay increased to about 500 days, or nearly one and a half years. In 1919, deliveries improved, but the delay remained lengthy, close to 150 days. Between 1921 and 1923, the delay was still 100 days. By 1927, the journal was delivered almost as quickly as in the prewar period. The pattern for the *Annalen der Physik*, the German journal that published Albert Einstein's famous 1905 papers, looks similar (Figure I, Panel A).

We also plot delays for two Allied journals from abroad, the French journal *Comptes Rendus* and the British journal *Nature*, the leading general scientific journals from these countries. Before the war, the *Comptes Rendus* arrived about 21 days after publication (Figure I, Panel B). By 1917, the delay increased to about 45 days. By 1919, the delay extended to 57 days, about three times longer than in the prewar period. After 1921, the delay returned to its prewar level. Before the war, *Nature* arrived only 10 days after publication—faster than the other journals, presumably because of shorter shipping routes from Britain. The delay for *Nature* almost tripled to 27 days during the war, and then partly recovered to about 19 days by 1921.

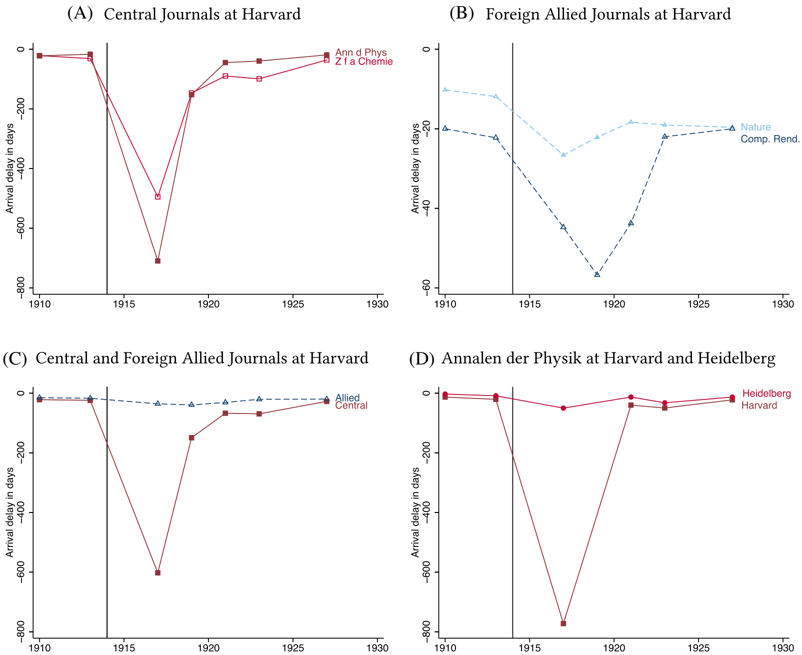

FIGURE I

Arrival Delay of International Journals

Panel A plots the average delay between publication and arrival date at the Harvard library for the German journals *Zeitschrift für Analytische Chemie* and *Annalen der Physik*. Arrival dates are based on library entry stamps (see Online Appendix Figure A.1 for an example). Delays are calculated as yearly averages for 1910, 1913, 1917, 1919, 1921, 1923, and 1927. Panel B plots the delay for two Allied journals, the British journal *Nature* and the French journal *Comptes Rendus*. Panel C compares average delays for German journals and Allied journals. Panel D compares delays for the *Annalen der Physik* at Harvard and at the University of Heidelberg in Germany. In Panel D, the delay at Harvard differs slightly from the delay reported in Panel A because we focus on journal issues that were available both at Harvard and at Heidelberg. Data on entry stamps were collected by the authors at Harvard and at the University of Heidelberg (see Online Appendix E.1 for details).

While the arrival delay for all foreign journals increased during the war and the boycott, the delay for German journals increased markedly more than for Allied journals (Figure I, Panel C). To investigate whether the increase in arrival delays for German journals was caused by a general disruption of the German publishers, we compare arrival delays for the *Annalen der Physik* at Harvard and the German University of Heidelberg (see Online Appendix E.1 for details). Even at Heidelberg, the delay increased during the war, but nowhere near as much as at Harvard (Figure I, Panel D).

These patterns indicate that foreign journals, particularly those from the enemy camp, became harder to access during the war and the boycott. Moreover, since Harvard has one of the best-funded university libraries in the world, it is plausible that the delays experienced by other universities were more extensive.

2. Scientific Conferences. The war and the boycott also impacted international scientific conferences. Between 1900 and 1914, scientists held 443 large international congresses. Between 1915 and 1918, only seven international congresses took place (Forschungen und Fortschritte 1933). In the postwar period, the number of international congresses was less than 20 in 1919, but steadily increased to 110 in 1926, and to 165 in 1930 (Kerkhof 1940).[6] During the boycott, Central scientists were banned from most international conferences. While this ban was strictly enforced in the first postwar years, it continued to limit conference attendance of Central scientists until 1926. Kerkhof (1940) reports that the ban on German scientists applied to all international conferences in 1919; to about 85% in 1920; to about 60% in 1921 and 1922; and to about 50% in 1924 and 1925. After 1926, German scientists were excluded from fewer than 15% of international conferences.

We complement the historical accounts with data on attendance records of the International Congress of Mathematicians (ICM), the largest mathematics conference. In the prewar period, Germany always sent large delegations to the ICM (see Table II). The 1916 congress that was scheduled to be held in Stockholm was canceled because of the war. The first postwar congress in 1920 was not held in Stockholm but was relocated to Strasbourg in a symbolic move. Strasbourg lies in the Alsace region that had been annexed by Germany in the 1870–1871 war with France and was returned after WWI. German mathematicians were neither invited to Strasbourg (1920 congress) nor to Toronto (1924 congress). By 1928, the boycott had ended, and Germany sent the second largest delegation, after the host nation, to Bologna.

6. These figures only refer to large international congresses, such as the International Congresses of Mathematicians below, and not to smaller international workshops. We are not aware of systematic data for the smaller gatherings.

TABLE II

ATTENDANCE OF INTERNATIONAL CONGRESSES OF MATHEMATICIANS

Year	Location	Delegates from:							
		Germany	Switzerland	France	U.S.	Canada	U.K.	Italy	Others
1897	Zürich	53	68	29	7	0	3	25	57
1900	Paris	26	7	93	19	1	12	23	69
1904	Heidelberg	204	13	29	19	1	8	14	108
1908	Rome	174	18	92	27	1	33	213	142
1912	Cambridge (U.K.)	70	10	45	87	5	270	41	181
1916	Stockholm	Canceled							
1920	Strasbourg	0	12	112	15	1	11	7	99
1924	Toronto	0	5	45	270	118	93	15	80
1928	Bologna	106	48	91	76	7	64	412	312
1932	Zürich	142	185	89	102	2	49	81	203

Notes. The Table reports the number of delegates at each *International Congress of Mathematicians*. Data were collected by the authors from historical issues of *Proceedings of the International Congresses of Mathematicians* (see Online Appendix E.2 for details).

We further document that even small and very elitist conferences were affected by the war and the boycott. We analyze attendance patterns at the Solvay Conferences in Physics. Nobel laureate Werner Heisenberg lauded "the Solvay meetings . . . as an example of how much well planned and well organized conferences can contribute to the progress of science" (Mehra 1975, VII). The first Solvay Conference was organized in 1911 and was attended by the leading physicists of the time, including Marie Curie, Ernest Rutherford, Max Planck, and Albert Einstein (Figure II, Panel A and Online Appendix Table A.2). In that year, 9 of the 24 participants came from Central countries. In 1913, 9 of the 31 participants came from Central countries. During the war, the Solvay Conferences were discontinued. The first postwar conference took place in 1921. Scientists from Central countries were not invited.[7] Nor were they invited to the 1924 conference. By 1927, the boycott had ended and 5 of the 30 participants came from Central countries.[8] The 1927 conference is possibly the most

7. The lone German invited to the 1921 and 1924 conferences was Albert Einstein, then a professor at the University of Berlin. The invitations reflected his special status in the scientific community and his reputation as an avid internationalist. He declined to attend in 1921 for personal reasons and in 1924 because none of his German colleagues had been invited (Mehra 1975, XXIII).

8. Two more participants were de facto in the German system but are classified as Neutrals in Mehra's data and hence not circled in Figure II, Panel E. Heisenberg had a joint appointment at the German University of Göttingen and the Danish University of Copenhagen and accepted a professorship at the German University of Leipzig in 1927. Schrödinger moved to the University of Berlin in 1927.

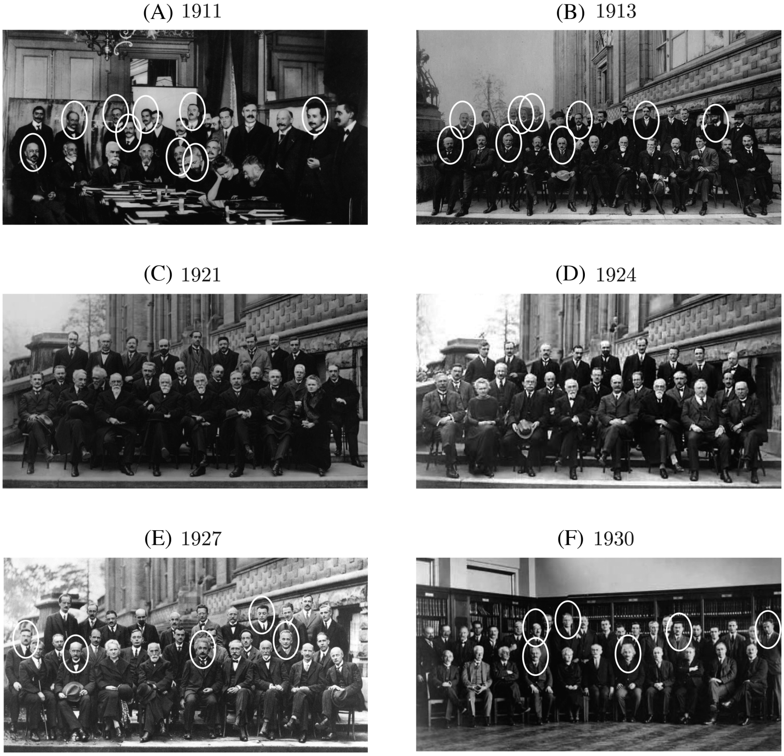

FIGURE II

Central Attendance at Solvay Conference

The figure shows delegates at the Solvay Conferences in physics. Circles indicate delegates from Central countries. See Online Appendix Table A.2 for delegate names. Data were collected by the authors from Mehra (1975) (see Online Appendix E.3 for details).

famous scientific conference ever organized. It took place at the height of the quantum revolution, and 17 of the 30 participants were current or future Nobel laureates. In 1930, 6 of the 36 participants came from Central countries.

III. DATA

III.A. *Censuses of University Scientists for 1900 and 1914*

We obtain data from various sources. First, we collect two historical censuses of all university scientists in the world for

1900 and 1914. The data come from two volumes of *Minerva—Handbuch der Gelehrten Welt*, the most comprehensive worldwide listing of university professors for this period. We digitize more than 2,500 pages that list university professors of all ranks (e.g., assistant, associate, and full professors), of all fields, and from all universities in the world (see Online Appendix Figure A.2 for a sample page).

The data contain information on 569 universities in the year 1900 and 973 universities in the year 1914 (Online Appendix Table A.1, Panel A). Across all fields, we manually digitize the names, affiliations, and fields of 23,917 professors in 1900 and 36,777 professors in 1914 (Online Appendix Table A.1, Panel A). Online Appendix Figure A.3 shows the distribution of scientists in 1914. The map illustrates the concentration of scientific activity in the United States and Western Europe.

We focus our empirical analysis on five scientific fields: medicine, biology, chemistry, physics, and mathematics. We concentrate on these because at that time scientists in these fields already established a habit of publishing the majority of their research in scientific journals.[9] Our data contain information on 10,133 scientists in 1900 and 15,891 scientists in 1914 across the five fields (Online Appendix Table A.1, Panel B).

III.B. Publication and Citation Data

We also collect all papers that were published in 160 top scientific journals from the *ISI Web of Science* for the period 1900 to 1930 (see Online Appendix E.4 for details on the selection of journals and Online Appendix Table A.3 for a list of the 160 journals), including information on the cited references (see Online Appendix E.4 for a detailed description of how we obtain the full list of authors and citations for all cited references).

9. *Minerva* lists the exact specialization of each scientist. Many mathematicians, for example, do not report "mathematics" but "algebra" or "analysis," often in native languages, as their field. We manually recode several thousands of the exact specializations into 32 fields (biology, physics, history, law, and so on).

The publishing process closely resembled publishing in modern times.[10]

The analysis crucially depends on knowing the country of authors and cited references. Most historical scientific journals, however, did not report author affiliations. For example, Max Planck's famous 1901 paper "On the Law of Distribution of Energy in the Normal Spectrum," which laid the foundation for the quantum revolution, did not include Planck's affiliation.

We assign countries to authors and references in a three-step, hierarchical process (see Online Appendix E.4 for further details). First, we use the country information from the affiliation reported in those papers that list affiliations. Second, we use the country information from the two scientist censuses.[11] Third, we expand the country information for authors with identical names within the corresponding citing or cited journal. Consider the example of Nobel laureate Arthur Compton. "A. Compton" published a paper in the *Physical Review* in 1923 with a U.S. affiliation and another paper in the same journal in 1920. Because the 1920 paper did not report an affiliation, we use the affiliation information from the 1923 paper to assign a U.S. affiliation to the 1920 paper.

We assign countries to papers and references using the fraction of citing authors and referenced authors from each country. A paper (or reference) exclusively written by authors from the United States, for example, counts as one U.S. paper. A paper coauthored by one U.S. author and one U.K. author counts as 0.5 U.S. paper and 0.5 U.K. paper.[12]

10. Because the historical part of the *Web of Science* focuses on the highest-cited journals, it has very good coverage of Anglo-Saxon and German journals. The coverage of French journals, for example, is less comprehensive. This does not bias our analysis because our regressions implicitly control for persistent differences in coverage across countries.

11. In the very rare cases that two or more scientists had identical names and worked in the same field but in different countries, we assign the paper proportionally to each country. For example, the censuses contain two chemists with the name J. Schmidt, one in Germany and one in Austria. We therefore count chemistry papers published by J. Schmidt as half German and half Austrian. Note that the *Web of Science* only reports the last name and initials of each author.

12. The country of papers and scientists is assigned using the scientist's university affiliation. Between 1900 and 1914, 2.75% of scientists in Allied and Central countries moved across countries and 1.11% moved across camps. For papers in journals that report affiliations, the moves are reflected in our data. For papers that do not report affiliations, moves after 1914 will not be recorded and authors will remain assigned to the country where they worked in 1914.

Mistaking an author for another author with the same name from the same country does not introduce measurement error because the sole purpose of this matching is the assignment of countries to citing authors and referenced authors. Remaining mistakes in assigning countries to papers and references will introduce measurement error. Depending on the estimated specification, the measurement error will either affect the dependent variable or the explanatory variables. With classical measurement error, our results remain unbiased in the first case and will be biased toward zero in the second case. The latter would make it more difficult to find significant effects.

III.C. Data on Nobel Prize Nominations

To measure scientific breakthroughs, we also collect data on nominations for the physics, chemistry, and physiology/medicine Nobel Prizes from Nobelprize.org (2014). The data contain 993 individuals who received at least one nomination for a Nobel Prize between 1905 and 1945. We merge these data with the publication data from the *Web of Science* to identify research that was worthy of a Nobel Prize nomination (see Section V.B.1 for details).

III.D. Full Text of U.S. Patents between 1920 and 1979

To assess how basic science produced in this period was applied in the development of new technology we obtain the full text of more than 2.5 million U.S. patents for the years 1920 to 1979 from the U.S. Patent Office (see Online Appendix E.4 for details). The 2.5 million patents contain more than 7.5 billion words. We then search these data for novel words that were introduced by scientific papers between 1905 and 1930.

III.E. Final Data Sets

We combine these data to construct two data sets: a paper-level data set that allows us to study changes in international citations and the similarity of papers (in Section IV) and a scientist-level data set that allows us to study how the breakdown of international scientific cooperation affected the productivity of scientists (in Section V).

The paper-level data set covers the period 1905 to 1930 and contains all papers for which we match the country of at least one author and at least one reference, and for which the

Web of Science reports the number of times the references are cited until today.

The scientist-level data set is a panel data set of all university scientists who published at least one paper between 1905 and 1930. It contains yearly productivity measures for each scientist.

IV. International Citations and the Similarity of Papers

We use the paper-level data to quantify how WWI and the boycott impacted references in scientific papers and the similarity of paper titles. These measures are attempts to proxy for international knowledge flows. Directly measuring knowledge flows between all scientists in the world would be nearly impossible. For example, one could not know or quantify whether scientists were aware of certain papers or whether they engaged in discussions about specific research topics in formal or informal scientific gatherings with their colleagues.

IV.A. The Effect of WWI and the Boycott on International Citations

First we measure changes to international citations in scientific papers. For each paper, we group references as follows: references to existing research from home, to foreign research from inside the camp, or to foreign research from outside the camp.[13] We divide these counts by the total number of references and obtain three shares: the share of citations to home ($\frac{c_{Home}}{C_{Total}}$), foreign countries inside the camp ($\frac{c_{Foreign-IN}}{C_{Total}}$), and foreign countries outside the camp ($\frac{c_{Foreign-OUT}}{C_{Total}}$).

To measure citations to recent research, we consider references to research published in the preceding five years.[14] The average paper in our sample includes 17.6 references overall; of these, 7.4 cite recent research, and 4.6 cite recent research published in one of the 160 journals in our data. Out of these 4.6 references, we are able to match the country to 3.0 references. For 2.6 of these references, the *Web of Science* reports the number of times the reference had been cited until today.

13. For the main results, we exclude self-citations when we count the references to research from home. The results are robust to including self-citations as citations to research from home (see Online Appendix Table A.7).

14. The results are robust to considering research published in the preceding 3 or 10 years as recent research (see Online Appendix Table A.8).

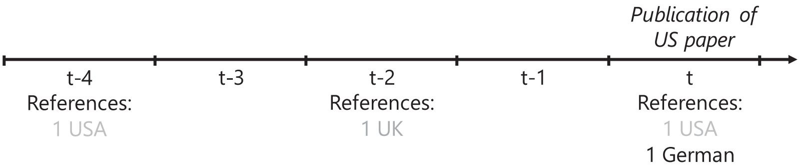

FIGURE III

Example Citing Paper and References

Figure III illustrates our measure. A paper published by a U.S. author in year t includes four references to research published in the preceding five years; one reference to U.S. research that was published in year t, one reference to German research that was published in year t, one reference to U.K. research that was published in year $t-2$, and one reference to U.S. research that was published in year $t-4$. The corresponding shares are:

$$\frac{c_{Home}}{C_{Total}} = \frac{2}{4} = 0.5, \quad \frac{c_{Foreign-IN}}{C_{Total}} = \frac{1}{4} = 0.25, \\ \text{and} \quad \frac{c_{Foreign-OUT}}{C_{Total}} = \frac{1}{4} = 0.25.$$

Table III summarizes the citation shares in our sample. About 69% of references quote research from home, 16% quote research from foreign authors inside the camp (e.g., U.S. papers quoting research from the United Kingdom), and about 15% quote research from outside the camp (e.g., U.S. papers quoting research from Germany). If we consider citations to the very best research, as measured by references that quote research that ended up in the top 1% of the citation distribution, 5.4% of references quote top research from home, 1.2% quote top research from foreign authors inside the camp, and about 1.3% quote top research from outside the camp.

1. Citations to All Research. We create three observations per paper: the share of references quoting research from home, from inside the camp, and from outside the camp. We then investigate how these shares (Y_{ic}) changed after 1914 by estimating the

TABLE III

SUMMARY STATISTICS: INTERNATIONAL CITATION SHARES AND LSA TITLE SIMILARITY

	(1)	Home (2)	Foreign inside camp (3)	Foreign outside camp (4)
Panel A: References				
Quality of references	Aver. num. of cits. to recent research	Average citation shares to recent research		
All references	2.593	0.686	0.159	0.150
Top 1% references	0.207	0.054	0.012	0.013
Top 3% references	0.479	0.126	0.027	0.029
Top 5% references	0.702	0.181	0.041	0.040
Panel B: Standarized LSA title similarity				
		Average LSA title similarity to recent papers		
Similarity to most similar title		0.376	−0.300	−0.076
Avg. similarity to 5 most similar titles		0.462	−0.399	−0.063

Notes. In Panel A, column (1), the table reports the number of references to recent papers (published between year t-4 and t, where t is the publication year of the citing paper). In Panel A, columns (2)-(4), the table reports citation shares to research from home, foreign countries inside the camp, and foreign countries outside the camp. For top x% references, citation shares are computed over all references (both high quality and low quality). In Panel B, columns (2)-(4), the table reports the standardized (i.e., mean 0 and standard deviation 1) LSA title similarity to papers by scientists from home, foreign countries inside the camp, and foreign countries outside the camp. In the penultimate row, LSA title similarity is computed as the similarity to the most similar title from each camp. In the last row, LSA title similarity is computed as the average similarity to the five most similar titles from each camp. We compute title similarity to recent papers (published between year t-4 and t). The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

following regression:

$$
\begin{aligned}
(1)\quad Y_{ic} = {} & \omega_1 \cdot 1\left[c = \text{Foreign Out}\right] \\
& + \omega_2 \cdot 1\left[c = \text{Foreign Out}\right] \times 1\left[t(i) = \text{Post-1914}\right] \\
& + \iota_1 \cdot 1\left[c = \text{Foreign In}\right] \\
& + \iota_2 \cdot 1\left[c = \text{Foreign In}\right] \times 1\left[t(i) = \text{Post-1914}\right] \\
& + \text{Citing Paper} FE_i + \epsilon_{ic},
\end{aligned}
$$

where i indexes citing papers and c indexes camps. A home indicator is excluded from the regression. Hence, ω_1 measures how the prewar share of references to research from outside the camp differed from the prewar share of references to research from home. Similarly, ι_1 measures how the prewar share of references to research produced by foreign authors from inside the camp differed from the prewar share from home. The parameters of interest, ω_2 and ι_2, measure how the foreign shares changed after 1914, relative to the home share.

The regression also includes a fixed effect for each citing paper. These fixed effects control for changes in citation patterns over time because the sum of all paper fixed effects within a year are collinear with a year fixed effect. Similarly, the fixed effects control for permanent differences in citation patterns across countries, for example, if U.S. authors generally include more references to research produced at home (e.g., a U.S. fixed effect would be collinear with the sum of paper fixed effects for all U.S. papers). The paper fixed effects also control for permanent differences in citation patterns across fields, for example, if chemists always cite more research produced at home because the chemical industry is differently specialized across countries. The fixed effects also control for permanent differences across fields in a certain country, for example, if U.S. chemists generally cite more research produced at home. To account for potential correlations of regression residuals in a certain field-country pair, for example, chemistry in the United States, we cluster standard errors at the field-country level.

After the onset of WWI, papers cited relatively less research from outside the camp. The share of references quoting research from outside the camp fell by 0.22, relative to the home share (Table IV, column (1), significant at the 1% level), a reduction of 85% relative to the prewar share of references quoting

TABLE IV
CHANGES IN INTERNATIONAL CITATIONS

Dependent variable: *Citation Shares to recent research*	(1)	(2)	(3)	(4)
Foreign outside camp × post-1914	−0.217***	−0.261***		
	(0.033)	(0.040)		
Foreign outside camp × WWI			−0.222***	−0.229***
			(0.025)	(0.034)
Foreign outside camp × boycott			−0.245***	−0.258***
			(0.034)	(0.052)
Foreign outside camp × post boycott			−0.194***	−0.213***
			(0.042)	(0.051)
Foreign inside camp × post-1914	−0.072*	−0.155***		
	(0.041)	(0.051)		
Foreign inside camp × WWI			−0.111***	−0.148***
			(0.040)	(0.045)
Foreign inside camp × boycott			−0.089**	−0.164***
			(0.042)	(0.057)
Foreign inside camp × post boycott			−0.048	−0.154**
			(0.048)	(0.059)
Paper fixed effects	Yes	Yes	Yes	Yes
Camp fixed effects	Yes	Yes	Yes	Yes
Foreign in/outside time trends		Yes		Yes
Observations	105,378	105,378	105,378	105,378
Number of citing papers	35,126	35,126	35,126	35,126
Within R-squared	0.334	0.335	0.335	0.335

Notes. Each column reports one set of parameter estimates of regression (1) for citing papers published between 1905 and 1930. The dependent variable measures citation shares to research by scientists from home, foreign countries inside the camp, and foreign countries outside the camp. We count citations to recent research, that is, research published in the preceding five years, for example, 1901–1905 for citing papers published in 1905, 1902–1906 for citing papers published in 1906, and so on. The reference/omitted category is the citation share to research from home. Standard errors are clustered at the country-times-field level. Significance levels: *** $p < .01$, ** $p < .05$, and * $p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

research from outside the camp. The share of references quoting research from foreign authors inside the camp fell by 0.07, relative to the home share (Table IV, column (1), significant at the 10% level), a reduction of 50% relative to the prewar share of references quoting research from inside the camp. The decline in the share of references quoting research from outside the camp was significantly larger than the relative decline in the share of references quoting research from inside the camp (p-value < .001). The results for both camps are slightly larger, in absolute magnitude, if we include camp-specific linear trends in the regression (Table IV, column (2)).

The estimated effect varies over time. The relative decline of the share of references quoting research from outside the camp was 0.22 during WWI, 0.25 during the boycott, and 0.19 in the postboycott period (Table IV, column (3), all significant at the 1% level).[15] The relative decline in the share of references quoting foreign research from inside the camp was 0.11 during WWI, 0.09 during the boycott, and 0.05 in the postboycott period (Table IV, column (3), only the first two are significant at the 1% and 5% level, respectively). The results are slightly larger if we control for camp-specific linear trends (Table IV, column (4)).

To get a better understanding of the timing of these changes, we estimate yearly coefficients:

$$\begin{aligned} Y_{ic} = & \sum_{\tau=1905}^{1930} \omega_\tau \cdot \mathbf{1}[c = \text{Foreign Out}] \times \mathbf{1}[t(i) = \tau] \\ & + \sum_{\tau=1905}^{1930} \iota_\tau \cdot \mathbf{1}[c = \text{Foreign In}] \times \mathbf{1}[t(i) = \tau] \\ & + \text{Citing Paper } FE_i + \epsilon_{ic}. \end{aligned} \tag{2}$$

A home indicator is excluded from the regression. Hence, ω_τ measures how the share of references to research from outside the camp differed from the share of references to research from home in year τ. Similarly, ι_τ measures how the share of references to research produced by foreign authors from inside the camp differed from the share of references from home. We plot the yearly coefficients in Figure IV. Even before WWI, papers contained fewer references to recent research from outside the camp, and even fewer references to foreign research from inside the camp, indicating a substantial home bias (Figure IV). After the onset of the war, relative citations to research from foreign authors declined sharply, particularly for citations to research from outside the camp. Relative citation shares to research from outside the camp

15. It is important to keep in mind that we analyze references produced in the preceding five years for these results. For a paper published in 1919, for example, we count references to research published between 1915 and 1919.

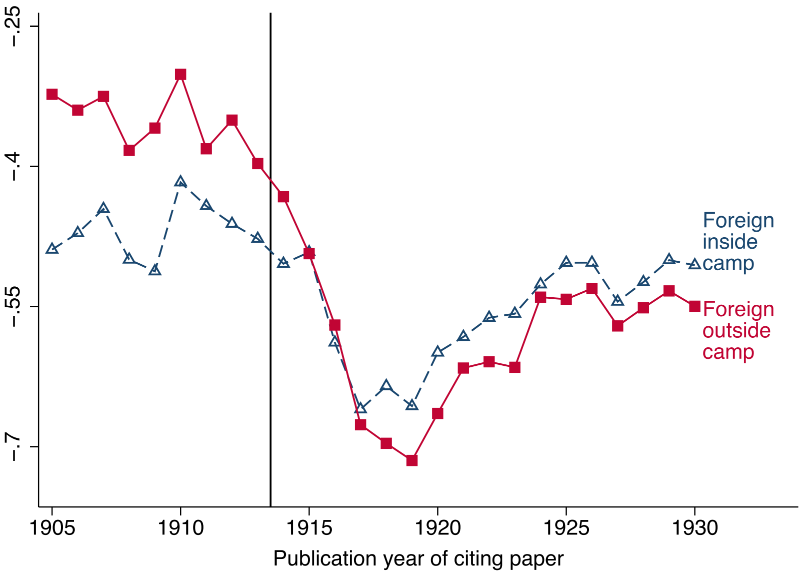

FIGURE IV

International Citation Shares Relative to Home

The figure plots parameter estimates of regression (2). The “Foreign outside camp” line reports point estimates (ω_τ) that measure citation shares to research from outside the camp, relative to research from home. The “Foreign inside camp” line reports point estimates (ι_τ) that measure citation shares to research from foreign scientists inside the camp, relative to research from home. We count citations to recent research, that is, research published in the preceding five years. For example, the first dot (1905) measures relative citation shares to research published between 1901 and 1905. The second dot (1906) measures relative citation shares to research published between 1902 and 1906, and so on. Point estimates and corresponding standard errors are reported in Online Appendix Table A.5. All point estimates are significantly different from 0 at the 1% level. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

declined from −0.35 before the war to about −0.71 at the end of the war and the early boycott, a decline of about 0.36. Relative citations shares to research from inside the camp declined from about −0.47 to −0.64, a decline of about 0.17. After 1919, citation shares began to recover but remained lower than in the prewar period.

2. Citations to Frontier Research. In further results, we explore whether citations to frontier research were also affected. We define the frontier as research that ended up in the top percentiles (top 5%, top 3%, and top 1%) of the field-level citation distribution. We count the total number of citations of each piece of research until today, that is, almost 100 years. This measure of the research frontier therefore captures the very long-run view of the quality of research and it is less likely to be affected by short-term scientific "fashions."[16]

The share of references to top 5% research from outside the camp fell by 0.053, relative to references to top 5% research from home (Table V, column (1), significant at the 1% level), a reduction of 95%, relative to the prewar share. By construction, the share of references that quote top 5% research is smaller than the share of references that quote research of any quality (see Table III), and hence, coefficients are likely to be smaller in absolute terms. However, in percentage terms the relative declines were similar. The point estimate becomes larger in absolute magnitude if we control for linear camp-specific trends (Table V, column (2)).

The share of references to top 5% research from foreign authors inside the camp fell by 0.023 relative to top 5% research from home (Table V, column (1)), a reduction of 72% relative to the prewar share. The relative decline in the share of references to top 5% research from outside the camp was significantly larger than the relative decline in the share of references to top 5% research from inside the camp (p-value $< .001$). Yearly coefficients are reported in Figure V, Panel A.

We also find that the share of references to top 3% or top 1% research from outside the camp fell significantly, with percentage declines of 95% and 131%, respectively (Table V, columns (3)–(6), significant at 1%, also Figure V). The share of references to top 3% or top 1% research from foreign authors inside the camp also fell, but by less than the share of references to research from

16. Specifically, we divide the share of references to research from home into references that ended up in the top 5% of the distribution and references that ended up in the bottom 95%. Similarly, we divide the shares to research from inside the camp and outside the camp. Hence, the data now contain six observations per paper. Citations to top research from home are the omitted category. The top 5% is measured at the subject level for all papers in the 160 journals in our data, independently of whether we can assign countries to authors and/or references. We construct analogous measures of citations to research that ended up in the top 3% or top 1% of the citation distribution.

TABLE V

CHANGES IN INTERNATIONAL CITATIONS: FRONTIER RESEARCH

Dependent variable:	Frontier: 5%		Frontier: 3%		Frontier: 1%	
Cit. Sh. to recent frontier research	(1)	(2)	(3)	(4)	(5)	(6)
Foreign outside camp × post-1914	−0.053*** (0.017)	−0.097*** (0.021)	−0.035*** (0.013)	−0.066*** (0.013)	−0.021*** (0.006)	−0.039*** (0.007)
Foreign inside camp × post-1914	−0.023 (0.015)	−0.071*** (0.021)	−0.019* (0.011)	−0.049*** (0.013)	−0.013** (0.006)	−0.033*** (0.007)
Paper fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Camp fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Nonfrontier research interactions	Yes	Yes	Yes	Yes	Yes	Yes
Foreign *in*/*outside* time trends		Yes		Yes		Yes
Observations	210,756	210,756	210,756	210,756	210,756	210,756
Number of citing papers	35,126	35,126	35,126	35,126	35,126	35,126
Within R-squared	0.235	0.235	0.299	0.300	0.400	0.400

Notes. Each column reports one set of parameter estimates of regression (1) for citing papers published between 1905 and 1930. The dependent variable measures citation shares to frontier and nonfrontier research by scientists from home, foreign countries inside the camp, and foreign countries outside the camp, that is, six shares for each citing paper. The table only reports estimates for frontier research, although the regressions control for nonfrontier times post-1914 indicators. For the results reported in columns (1)–(2), frontier research is defined as research that ended up in the top 5% of the subject-level citation distribution, counting citations until today. Similarly, for the results reported in columns (3)–(4) (and (5)–(6)), frontier research is defined as research that ended up in the top 3% (and 1%) of the subject-level citation distribution. We count citations to recent research, that is, research published in the preceding five years: e.g., 1901–1905 for citing papers published in 1905, 1902–1906 for citing papers published in 1906, and so on. The reference/omitted category is the citation share to frontier research from home. Standard errors are clustered at the country-times-field level. Significance levels: $^{***}p < .01$, $^{**}p < .05$, and $^{*}p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

outside the camp (Table V, columns (3)–(6), significant at 1%, also Figure V). These results indicate that the war and the boycott not only affected citations to average research but also had significant and large effects on citations to high-quality research.

3. *Robustness.* It is important to note that potential changes in relative quality of scientific output in the Allied or Central camp are unlikely to explain our findings, because such changes would have decreased the share of references to research from outside the camp for one of the camps, but would have increased the share for the other camp.

The results are robust to a number of alternative specifications: to restricting the sample of citing papers to papers by authors with a university position by 1914, to only measuring citations to research published by authors with a university position by 1914, and to normalizing citation shares by the total number of potentially citeable papers produced in each camp. We

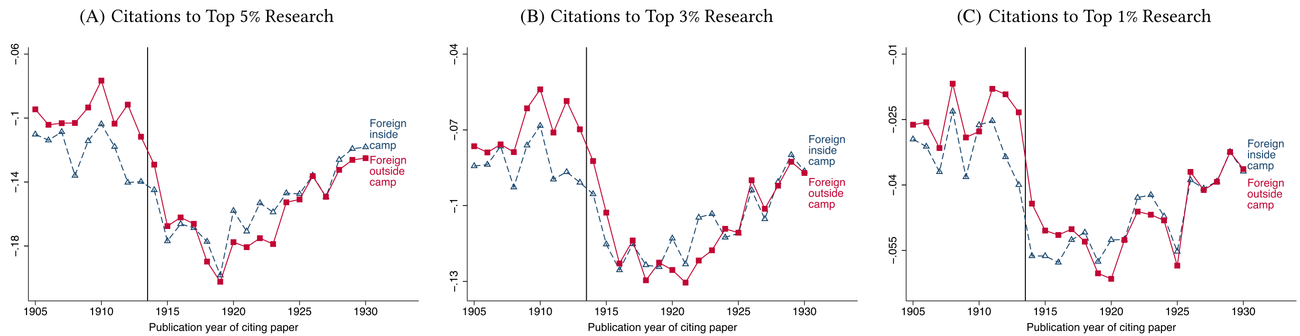

FIGURE V

International Citation Shares Relative to Home: Citations to High-Quality Research

Each panel plots one set of parameter estimates of a variant of regression (2). For the results reported in Panel A, we split citation shares to research from home into research that ended up in the top 5% of the citation distribution and research that ended up in the bottom 95%. Similarly, we split citation shares to research produced inside the camp and outside the camp. The “Foreign outside camp” line measures citation shares to top 5% research from outside the camp, relative to top 5% research from home. The “Foreign inside camp” line measures citation shares to top 5% research from foreign scientists inside the camp, relative to top 5% research from home. The regressions also include the citation shares to nonfrontier research from outside the camp, inside the camp, and home. In all panels, we count citations to recent research, that is, research published in the preceding five years. For example, the first dot (1905) measures relative citation shares to top 5% research published between 1901 and 1905, and so on. Panel B reports estimates for changes in citation shares to research that ended up in the top 3% of the citation distribution, and Panel C reports estimates for changes in citation shares to research that ended up in the top 1% of the citation distribution. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

also show that results are somewhat stronger for Allied than for Central scientists (see Online Appendix B.1 for details).

We also find that citation patterns in Neutral papers look quite different. Citations toward foreign research outside the Neutral camp do not decline during WWI or the boycott (see Online Appendix Figure A.8 and Online Appendix B.3).

IV.B. Do Changes in Citations Reflect Changes in International Knowledge Flows?

The observed changes in citations could reflect reduced international knowledge flows, that is, scientists not being aware of foreign research. Alternatively, the changes in citations could be a result of political hostility, that is, scientists knowing of foreign research but deliberately deciding not to cite it. If the effect were predominately driven by political hostility, presumably scientists also would have reduced citations to prewar research. We investigate reductions in citations to prewar research by investigating two cohorts of research (1903–1905 and 1911–1913).[17]

We find no evidence for a large dip in citations to prewar research from foreign countries during WWI or the boycott (Figure VI). The share of references quoting prewar research from foreign countries increased over time relative to the share of references quoting prewar research from home (the excluded category), because, even under normal conditions, knowledge takes time to reach foreign countries. Over time, citations to less-relevant work from home fade, but good papers from all camps continue to receive citations. We also estimate a variant of equation (1) for all prewar cohorts between 1903 and 1913. Citations to foreign research do not decline after 1914 for any of the nine prewar cohorts (see Online Appendix Table A.9).

These auxiliary results suggest that the changes to citations of recent research (Section IV.A) were predominantly driven by scientists' lack of knowledge about recent foreign research and not by political hostility. To further probe the effect of political hostility on citation shares, we also investigate citations to recent

17. These results fix the cohort of research (either to 1903–1905 or to 1911–1913) and investigate how citation shares to those two cohorts changed over time. In contrast, the main citation results investigate citation shares to a moving window of references, that is, references to research published between 1901 and 1905 for citing papers published in 1905, but to research published between 1902 and 1906 for citing papers published in 1906, and so on.

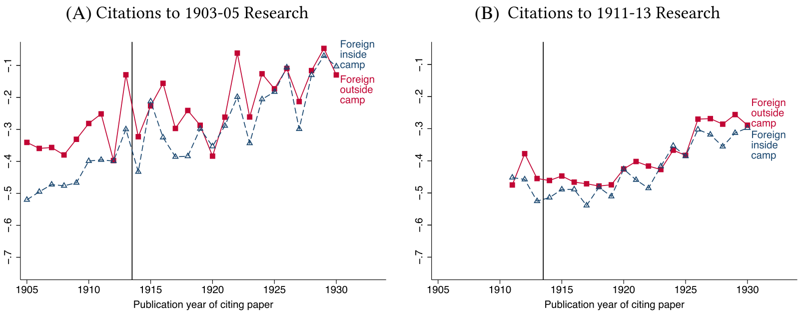

FIGURE VI

International Citation Shares Relative to Home: Citations to Prewar Research

Each panel plots one set of parameter estimates of regression (2) with citation shares to prewar research as the dependent variable. Differently from previous figures, each dot of any line measures relative citation shares to a fixed cohort of research, published either in 1903–1905 (Panel A) or in 1911–1913 (Panel B). In Panel A, the "Foreign outside camp" line reports point estimates (ω_τ) that measure citation shares to 1903–1905 research from outside the camp, relative to 1903–1905 research from home. The "Foreign inside camp" line reports point estimates (ι_τ) that measure citation shares to 1903–1905 research from inside the camp, relative to 1903–1905 research from home. Panel B refers to research published in 1911–1913 and can be interpreted analogously. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

research for nonchemists. The involvement of prominent chemists in the development of chemical weapons led to particularly strong resentment of chemists in the opposing camp. If political resentment were the main driver of changes to citations to recent research, we would expect smaller changes to citations shares if we excluded chemists from the sample. However, the results do not change substantially if we omit chemists (see Online Appendix Figure A.7b).

IV.C. The Effect of WWI and the Boycott on the Similarity of Paper Titles

1. Using Latent Semantic Analysis to Measure the Similarity of Titles. To complement our citation analysis, we investigate how WWI and the boycott affected the similarity of papers produced in the different camps. We analyze the similarity of papers by applying latent semantic analysis (LSA) (Deerwester et al. 1990; Landauer, Foltz, and Laham 1998) to the titles of scientific papers. LSA is a machine learning technique that retrieves semantic connections between words, so that even titles with completely different words can be classified as similar if the words are regularly used in similar contexts. For example, "n-dimensional," "manifold," and "topology" often appear together in paper titles. Therefore, LSA will classify a title that only contains "manifold" as similar to a title that only contains "topology." Moreover, LSA recognizes when the same word is used in different contexts. Thus, LSA offers a significant improvement over measures of similarity based solely on word counts.[18]

Because ISI translated all titles into English, we do not have to consider differences in original publishing languages when applying LSA. We prepare the titles for LSA by removing stopwords and one-letter words. We then use a Snowball stemmer to reduce the words to their morphological roots, so-called stems (Porter 1980, 2001). Finally, we remove titles with fewer than five stems, because titles with very few stems may have artificially high similarity. This leaves us with 79,438 paper titles D and a vocabulary V of 35,119 unique word stems, that is, terms, which we use to create a $D \times V$ document-term matrix.[19] The individual word counts in the matrix are then reweighted by their term frequency-inverse document frequency. This reweighting decreases the relative importance of words that carry little information but appear in many documents, for example, "study."

LSA uses truncated singular value decomposition to reduce the dimensionality of the document-term matrix from $D \times V$ to a user-chosen number of components C (for a detailed explanation see Online Appendix C). The output of LSA is a $D \times C$ document-component matrix with rows δ_d of dimension $1 \times C$. The components capture the semantic relationships between the documents.

We then use the document-component matrix to measure the similarity of titles by calculating the cosine similarity, a standard similarity measure in machine learning. The cosine similarity of document-pair i and j is defined as $\frac{\sum_{c=1}^{C} \delta_{i,c}\delta_{j,c}}{\sqrt{\sum_c \delta_{i,c}^2}\sqrt{\sum_c \delta_{j,c}^2}}$, where $\delta_{i,c}$

18. LSA also outperforms other machine learning techniques such as latent Dirichlet allocation (Blei, Ng, and Jordan 2003) or nonnegative matrix factorization (Lee and Seung 2001) in word-similarity tasks (Stevens et al. 2012).

19. We apply LSA to all papers published between 1905 and 1930, independently of whether we know the country of authors and references, because a larger set of papers improves the accuracy of LSA. When we estimate how title similarity changed during the war and the boycott we have to limit the sample to papers where we know the country of authors and references.

and $\delta_{j,c}$ are the elements inside the document-component matrix for documents i and j. The cosine similarity is 1 for titles that are identical and 0 for titles that are completely different. For each paper published in year t, we calculate the following three measures with respect to papers published between years $t-4$ and t:

(i) the cosine similarity to the most similar paper from home (excluding papers by the same author),
(ii) the cosine similarity to the most similar paper from inside the camp,
(iii) the cosine similarity to the most similar paper from outside the camp.

We also calculate alternative similarity measures using the average cosine similarity for the five most similar papers from each camp. We standardize the similarity measures to have zero mean and unit variance (see Table III for average similarity to papers from each camp).

2. The Effect of WWI and the Boycott on Title Similarity. For each paper, we create three observations with the title similarity to papers from each of the three camps (foreign outside the camp, foreign inside the camp, and home). We estimate equation (1) with LSA title similarity as the dependent variable. The parameters of interest are ω_2 and ι_2, which measure how the title similarities changed after 1914, relative to the title similarity to papers produced at home.[20]

After 1914, the similarity to papers from outside the camp fell by 0.47 standard deviations, compared with the similarity to papers from home (Table VI, column (1), significant at the 1% level). This result is robust to controlling for camp-specific linear time trends (column (2)). The relative decline was 0.46 std. dev. during WWI, 0.53 std. dev. during the boycott, and 0.43 std. dev. in the postboycott years (column (3)). The relative similarity to papers from inside the camp, however, did not change significantly after 1914 (Table VI, columns (1)–(3)). The exception is a regression

20. The number of observations is smaller than for the citation share regressions because we focus on papers with titles that have at least five words after stemming and removing stopwords. We also drop papers for which we cannot compute the similarity to the home camp because for small countries our data sometimes only contain one home paper published between year $t-4$ and t.

TABLE VI
THE SIMILARITY OF PAPERS AS MEASURED BY LATENT SEMANTIC ANALYSIS

Dependent variable:	Most similar title				Average five most similar titles			
LSA title similarity to recent papers	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Foreign outside camp × post-1914	−0.467***	−0.557***			−0.608***	−0.678***		
	(0.095)	(0.126)			(0.131)	(0.165)		
Foreign outside camp × WWI			−0.459***	−0.517***			−0.575***	−0.639***
			(0.084)	(0.117)			(0.118)	(0.143)
Foreign outside camp × boycott			−0.530***	−0.648***			−0.678***	−0.806***
			(0.111)	(0.193)			(0.150)	(0.213)
Foreign outside camp × post boycott			−0.426***	−0.593***			−0.569***	−0.753***
			(0.099)	(0.188)			(0.134)	(0.192)
Foreign inside camp × post-1914	0.060	−0.164			0.058	−0.226		
	(0.151)	(0.140)			(0.191)	(0.178)		
Foreign inside camp × WWI			−0.019	−0.181			−0.054	−0.267*
			(0.137)	(0.117)			(0.177)	(0.141)
Foreign inside camp × boycott			0.006	−0.320**			0.005	−0.425***
			(0.161)	(0.152)			(0.202)	(0.160)
Foreign inside camp × post boycott			0.122	−0.343**			0.131	−0.482***
			(0.154)	(0.165)			(0.194)	(0.165)
Paper fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Camp fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Foreign in/outside time trends		Yes		Yes		Yes		Yes
Observations	71,586	71,586	71,586	71,586	71,586	71,586	71,586	71,586
Number of citing papers	23,862	23,862	23,862	23,862	23,862	23,862	23,862	23,862
Within R-squared	0.156	0.157	0.157	0.158	0.225	0.228	0.227	0.228

Notes. Each column reports one set of parameter estimates of regression (1) for papers published between 1905 and 1930. The dependent variable measures the standardized (i.e., mean 0 and standard deviation 1) LSA title similarity to papers by scientists from home, foreign countries inside the camp, and foreign countries outside the camp. In columns (1) to (4), LSA title similarity is computed as the similarity to the most similar paper from each camp. In columns (5) to (8), LSA title similarity is computed as the average similarity to the five most similar papers from each camp. We compute title similarity to recent papers, that is, papers published in the preceding five years: 1901–1905 for papers published in 1905, and so on. The reference/omitted category is the LSA title similarity to papers from home. Standard errors are clustered at the country-times-field level. Significance levels: $^{***}p < .1$, $^{**}p < .05$, and $^{*}p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

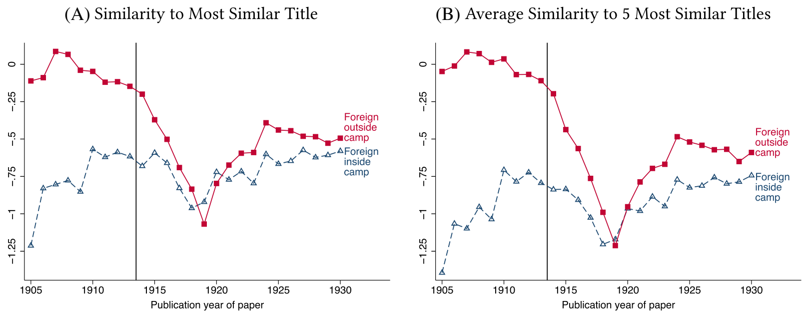

FIGURE VII

International Title Similarity Relative to Home

Each panel plots one set of parameter estimates of the equivalent of regression (2) where the dependent variable measures the standardized (i.e., mean 0 and standard deviation 1) LSA title similarity to papers by scientists from home, foreign countries inside the camp, and foreign countries outside the camp. In Panel A, LSA title similarity is computed as the similarity to the most similar title from each camp. In Panel B, LSA title similarity is computed as the average similarity to the five most similar titles from each camp. The "Foreign outside camp" line reports point estimates (ω_τ) that measure the LSA title similarity to papers from outside the camp, relative to papers from home. The "Foreign inside camp" line reports point estimates (ι_τ) that measure the LSA title similarity to papers from foreign scientists inside the camp, relative to papers from home. We measure title similarity to recent papers, that is, papers published in the preceding five years. For example, the first dot (1905) measures relative title similarity to papers published between 1901 and 1905, and so on. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

that controls for camp-specific linear time trends. According to this specification the similarity to papers from inside the camp fell by about 0.3 std. dev. during the boycott and the postboycott years (column (4)).

The results are very similar when we measure title similarity to the five most similar titles from each camp (Table VI, columns (5)–(8)). We also estimate yearly coefficients (Figure VII).[21] After

21. In most years, the similarity to the most similar paper from foreign countries outside the camp is larger than the similarity to the most similar paper from foreign countries inside the camp, because the data contain more papers from the United States and Germany. The probability of finding a similar paper is higher if a camp produces more papers. For Germany, foreign countries inside the camp are small (e.g., Austria) and hence it is less likely that we find similar papers produced by foreign scientists inside the camp. In contrast, foreign countries outside the camp are large (e.g., the United States) and hence it is more likely that we find similar papers. A similar argument applies to U.S. papers.

the beginning of WWI, the similarity to papers from outside the camp fell sharply, relative to the similarity to papers from home, and started to recover in the 1920s but did not reach its prewar levels until 1930. The relative similarity to papers from inside the camp also declined somewhat during the war, but less than the similarity to papers from outside the camp.[22] These results are robust to varying the number of components used to construct title similarity measures (Online Appendix Table A.11).

These findings corroborate the citation share results. It is important to note that LSA title similarity is exclusively computed from the information in paper titles of citing papers. In contrast, citation shares are computed from information in the references and do not use information from the titles of citing papers. While these results indicate that the scientific communities in enemy camps diverged during this period, this does not mean that this divergence was necessarily negative for scientific progress. We therefore investigate effects on scientific productivity in the next section.

We also investigate changes in title similarity of Neutral papers. The results for Neutral papers look quite different. During the war and the boycott, the similarity to papers from both outside the Neutral camp and foreign countries inside the Neutral camp did not change, relative to the similarity to papers from home (Online Appendix Figure A.10).

The temporary divergence of title similarity could either be caused by reduced international knowledge flows or by a war-related divergence of research motivated by military needs. For example, chemists in the opposing camps may have developed weapons relying on different scientific foundations. Excluding chemistry papers from the similarity analysis hardly changes the results (see Online Appendix Figure A.9), even though scientists in other fields were less involved in the war effort. This suggests that the divergence of research was at least partly driven by reduced international knowledge flows.

22. Because all titles are translated into English, the results are presumably not driven by diverging terminology in the two camps but rather by a divergence in the direction of research. Because LSA recognizes semantic context, it would also classify titles as similar if scientists in opposing camps temporarily used different scientific terms but later converged to a common terminology.

V. Interruption of International Cooperation and Scientific Productivity

V.A. Publications in Top Scientific Journals

Next we investigate whether the reduction of international scientific cooperation impacted scientific productivity. Because many scientists stress the importance of frontier knowledge, we compare productivity changes for scientists in fields that in the prewar period relied on frontier knowledge from abroad to changes for scientists in fields that relied on frontier knowledge from home. The productivity of scientists who relied on frontier knowledge from abroad, particularly from outside the camp, should be disproportionately affected by the breakdown of international scientific cooperation.

We proxy reliance on frontier knowledge from the three camps (home, foreign countries inside the camp, and outside the camp) with prewar citation shares, measured at the field-country level. Specifically, we compute citation shares in papers published by scientists in each field-country pair between 1905 and 1913. We also compute citation shares to nonfrontier research from the three camps. In Figure VIII, Panel A, we show how certain field-country pairs (e.g., chemistry in the United States) depended on research from home, foreign countries inside the camp, and outside the camp in the prewar period. In Panel B we show prewar dependence on frontier research.

A useful example of the identifying variation is the dependence on frontier research of biochemistry and biology in the United States. For U.S. scientists, knowledge from outside the camp came predominantly from Germany, while knowledge from foreign countries inside the camp came mainly from Britain. In biochemistry, Germany led the world in the early twentieth century. For example, the term "biochemistry" was coined by the German scientist Carl Neuberg in 1903. Biochemistry in Britain and the United States, however, had yet to take off. This was reflected in the prewar citation shares to frontier research of U.S. biochemists: 32% cited research from outside the camp, 12% from inside the camp, and 56% from home.[23] In biology, however,

23. Note that scientists in large countries across all fields disproportionately cited research from home. Our identifying variation relies on field-country-level differences in citations to research from home, foreign countries inside the camp, and outside the camp. Differences in the prewar reliance on foreign research between field-country pairs occur because field-country pairs produced different amounts of frontier research before the war (see Online Appendix D for a stylized example of the identifying variation).

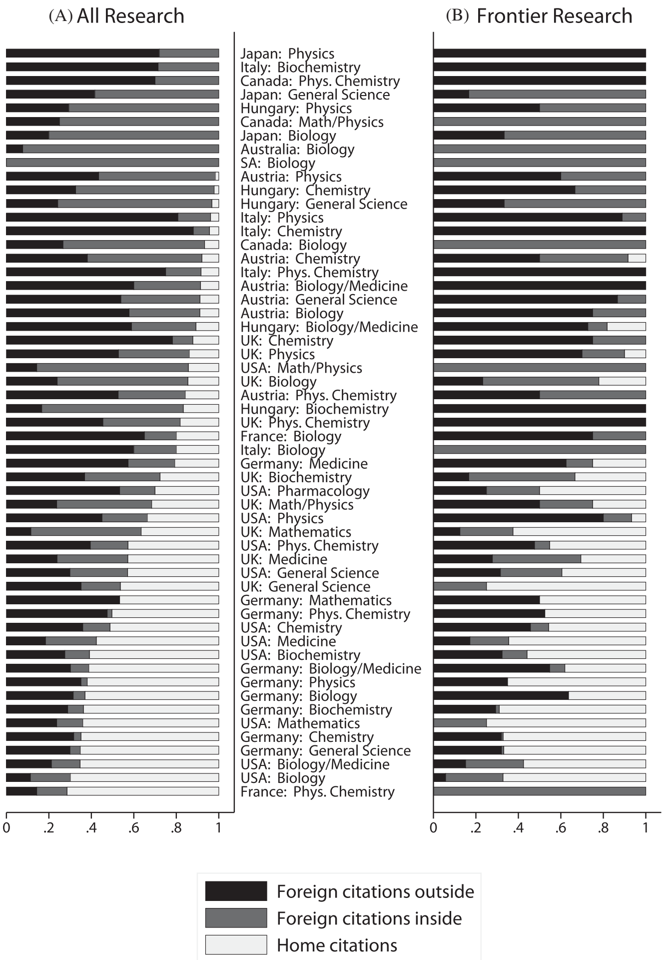

FIGURE VIII

Prewar Citations to Research from Home, Foreign Inside Camp, and Outside Camp

Panel A shows the prewar reliance on all research (i.e., both frontier and non-frontier) from home, from abroad outside the camp, and from abroad inside the camp for each field-country pair. Prewar reliance on all research is calculated as the average citation shares to recent research from home, foreign countries inside the camp, and foreign countries outside the camp for all citing papers published by all university scientists in each field-country pair between 1900 to 1913. Panel B focuses on prewar reliance on frontier research, measured as average shares of citations to top 3% research. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

Germany's influence was less pronounced, while Britain, and in particular the United States, contributed many important discoveries. This was reflected in prewar citation shares to frontier research of U.S. biologists: 6% cited research from outside the camp, 27% from inside the camp, and 67% from home.[24]

Average productivity of scientists in our sample declined during WWI and the boycott (see Online Appendix Figure A.11). We estimate the differential effect of the reduction in international scientific cooperation on productivity in a generalized difference-in-differences framework. We compare productivity changes of Allied and Central scientists in country-field pairs that relied on frontier research from foreign countries outside the camp and inside the camp, to productivity changes of scientists who relied on frontier research from home.[25]

$$\begin{aligned} Y_{ift} &= \beta_1 \cdot (\text{Prewar Reliance on Frontier OUT})_{if} \times 1[t = \text{Post-1914}] \\ &\quad + \beta_2 \cdot (\text{Prewar Reliance on Frontier IN})_{if} \times 1[t = \text{Post-1914}] \\ &\quad + \text{Scientist}FE_{if} + \text{Year}FE_t + X_{ift}\theta + \epsilon_{ift}, \end{aligned} \tag{3}$$

where i indexes scientists, f indexes field-country pairs, and t indexes years. For the first set of results, the dependent variable measures the number of publications per year for each scientist. The coefficients β_1 and β_2 measure productivity changes relative to scientists in field-country pairs that relied on frontier research from home (the excluded category). The regression includes a full set of scientist fixed effects that control for permanent differences in quality across scientists. The regression also includes a full

24. To simplify the exposition, our example and Figure VIII, Panel B focus on frontier research. However, in all regressions we use the prewar share of references quoting frontier research from home, nonfrontier research from home, frontier research from foreign authors inside the camp, nonfrontier research from inside the camp, frontier research from outside the camp, and nonfrontier research from outside the camp. For the main results presented in the paper, we require that scientists in a certain field-country pair published at least five papers before 1914 to construct the prewar dependence on research from home, foreign countries inside the camp, and foreign countries outside the camp. Choosing a higher threshold of prewar papers, for example, at least 10 papers, leads to very similar results (see Online Appendix Table A.13)

25. For scientists who worked in multiple fields, for example, physical chemistry and chemistry, we assign the reliance on frontier and nonfrontier research from the different camps according to the share of their publications in each field.

TABLE VII
SUMMARY STATISTICS: PRODUCTIVITY OF SCIENTISTS

	Mean	Std. dev.
Number of scientists	8,734	
Number of scientist-year observations	227,084	
Career age in years	7.444	7.708
Publications per year	0.267	0.950
Nobel-nominated papers per year	0.001	0.029
Nomination-weighted Nobel-nominated papers per year	0.003	0.152
Number of novel words (word innovation) per year	0.041	0.268
Patent-relevant word innovation per year	0.427	3.538

Notes. The table reports summary statistics for the panel of scientists with a university position by 1914. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt*, publication and citation data from *ISI Web of Science*, Nobel nomination and award data from Nobelprize.org (2014), and patent data from U.S. Patent Office (see Section III for details).

set of year fixed effects that control for yearly changes in productivity that affected all scientists in the same way, such as a reduction in productivity during the war years. We also control for the reliance on nonfrontier research from home, foreign countries inside the camp, and outside the camp, all interacted with post-1914 indicators. Furthermore, we control for five-year career-age indicators interacted with the main field of each scientist, that is, we control for different career-age productivity profiles for physicists, chemists, and so on. We estimate regression (3) for scientists who had a university position by 1914. This prevents potential selection bias caused by scientists of different quality entering or exiting the sample. The data contain 8,734 scientists with yearly productivity information for 1905 to 1930, which results in 227,084 scientist-year observations (Table VII). Standard errors are clustered at the country-times-field level.[26]

We estimate this regression for different definitions of the research frontier (top 1%, top 3%, or top 5%). Scientists in field-country pairs that relied on top 1% research from outside the camp published significantly less after 1914, compared to scientists who relied on top 1% research from home (Table VIII, column 1, significant at the 1% level). The estimated effect implies that scientists in a field-country pair that in the prewar period cited a lot of

26. We assign each scientist to his main research field according to his publications in each field.

TABLE VIII
EFFECT ON PUBLICATIONS

Dependent variable: *Number of publications*	Frontier: 1% (1)	Frontier: 3% (2)	Frontier: 5% (3)
Prewar reliance on frontier OUT × post-1914	−1.727***	−0.784***	−0.380*
	(0.638)	(0.282)	(0.220)
Prewar reliance on frontier IN × post-1914	−0.827	−0.363	−0.152
	(0.736)	(0.283)	(0.218)
Scientist fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Prewar reliance on nonfrontier	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes
Observations	227,084	227,084	227,084
Number of scientists	8,734	8,734	8,734
Within R-squared	0.062	0.062	0.062

Notes. Each column reports one set of parameter estimates of regression (3) for the panel of university scientists. The dependent variable measures the yearly number of publications in the 160 top journals in our data for the years 1905 to 1930. The number of publications is normalized by the number of authors and standardized to mean zero and standard deviation one within fields. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: $^{***}p < .01$, $^{**}p < .05$, and $^{*}p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

frontier research from outside the camp, such as biochemistry in the United States, published 0.1 of a std. dev. fewer papers per year after 1914 (i.e., 0.15 fewer biochemistry papers per year, a reduction of 33%), compared to scientists in field-country pairs that cited a lot of frontier research from home, such as U.S. biology. A field-country pair with one of the highest prewar reliance on frontier research from outside the camp was physics in Italy. Compared to a field-country pair that cited only frontier research from home, the estimated coefficient implies that Italian physicists published 0.27 std. dev. fewer papers per year after 1914 (i.e., 0.28 fewer physics papers per year, a reduction of 55%).

The productivity of scientists in field-country pairs that, in the prewar period, cited a lot of top 1% research from inside the camp also published less after 1914, but not significantly so (Table VIII, column (1)). The point estimate suggests that the relative productivity decline for scientists reliant on frontier research from inside the camp was about half as large as the productivity decline for scientists reliant on frontier research from outside the camp. This lines up well with the impact of WWI and the boycott on citation shares to frontier research from inside and outside the camp that we have shown in the first part of the paper (e.g., Table V, columns (1), (3), and (5)).

To understand the timing of these effects, we estimate yearly coefficients:

$$
\begin{aligned}
Y_{ift} = & \sum_{\substack{\tau=1905 \\ (\tau \neq 1913)}}^{1930} \beta_{1\tau} \cdot (\text{Prewar Reliance on Frontier OUT})_{if} \times 1[t=\tau] \\
& + \sum_{\substack{\tau=1905 \\ (\tau \neq 1913)}}^{1930} \beta_{2\tau} \cdot (\text{Prewar Reliance on Frontier IN})_{if} \times 1[t=\tau] \\
& + \text{Scientist}FE_{if} + \text{Year}FE_t + X_{ift}\theta + \epsilon_{ift}.
\end{aligned} \tag{4}
$$

Scientists in field-country pairs that relied on frontier research (as measured by the top 1%) from outside the camp suffered a sharp decline in productivity after 1914, compared to scientists who relied on frontier research from home (Figure IX). For these scientists, relative productivity did not recover. Scientists in field-country pairs that relied on frontier research from inside the camp suffered a smaller decline in productivity after 1914, which was not persistent. The figure also indicates that pretrends cannot explain the results. Scientists in field-country pairs that relied on frontier research from outside the camp, relative to frontier research from home, improved in the years until 1913, suggesting that, if anything, we underestimate the effect of the war and the boycott.[27] After 1913, however, the productivity of scientists in field-country pairs that relied on frontier knowledge from outside the camp declined sharply. Similarly, the figure indicates that pretrends cannot explain the productivity decline of scientists in field-country pairs that relied on frontier research from inside the

27. We report yearly coefficients and standard errors in Online Appendix Table A.12.

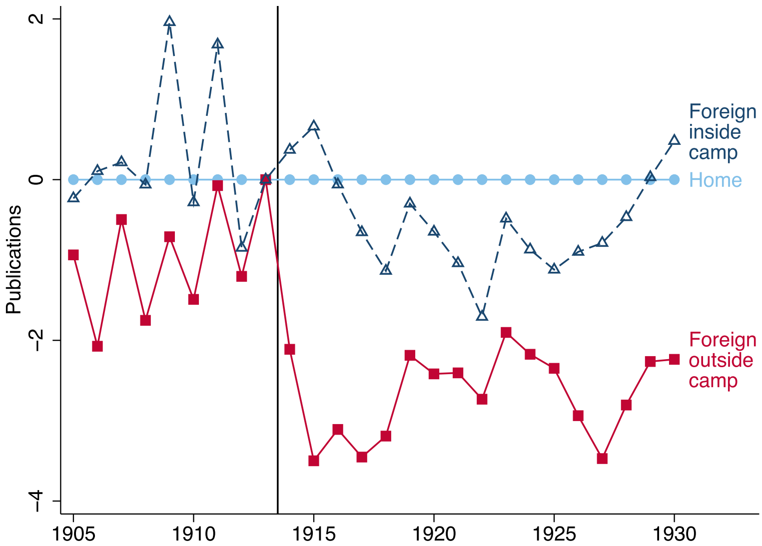

FIGURE IX

Effect on Publications

The figure plots parameter estimates from regression (4). The "Foreign outside camp" line reports point estimates ($\beta_{1\tau}$) that measure changes in yearly publications for scientists in field-country pairs that, in the prewar period, relied on frontier research from outside the camp, compared to scientists who relied on frontier research from home. The "Foreign inside camp" line reports point estimates ($\beta_{2\tau}$) that measure changes in yearly publications for scientists in field-country pairs that, in the prewar period, relied on frontier research from foreign scientists inside the camp, compared to scientists who relied on frontier research from home. Prewar reliance on frontier research is measured by prewar citations to frontier research at the field-country pair level. Frontier research is defined as research that ended up in the top 1% of the subject-level citation distribution, counting citations until today. The regression also controls for prewar reliance on nonfrontier research from each camp interacted with year indicators. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

camp, relative to scientists in field-country pairs that relied on frontier research from home.

If we alternatively measure the research frontier with top 3% research, we estimate a smaller, but still highly significant, productivity decline for scientists who relied on frontier research from outside the camp, compared to scientists who relied on frontier research from home. If we measure the frontier with top 5%

research, we estimate an even smaller, but still significant, decline in productivity (Table VIII, columns (2) and (3) significant at the 1% and 10% level, respectively). Scientists who relied on frontier research from foreign countries inside the camp suffered smaller and insignificant productivity declines. We test whether the productivity decline after 1914 was significantly larger for scientists reliant on frontier research from outside the camp than for scientists reliant on frontier research from inside the camp. For the different definitions of the knowledge frontier we get the following p-values: .136 (1% frontier), .055 (3% frontier), and .124 (5% frontier). For the most stringent specification with camp-times-field-times-year fixed effects (see Table IX) the p-values are: .147 (1% frontier), .007 (3% frontier), and .012 (5% frontier).

The alternative measures of frontier research suggest that the knowledge frontier is narrow-edged. Scientists who lost access to top 1% research experienced productivity declines that were about twice as large as the productivity declines of scientists who lost access to top 3% or top 5% research.[28]

1. Including Additional Fixed Effects. For the previous results, we normalize the dependent variable by the number of authors per paper. Without this normalization, the results remain very similar (Table IX, column (2)). Furthermore, the results are qualitatively unchanged if we include additional fixed effects (Table IX, columns (3)–(5)). These fixed effects control for various potential confounders that may be correlated with the propensity of citing frontier research from outside the camp, inside the camp, or home in the prewar period. In particular, we control for camp-times-year fixed effects to allow for cross-camp differences in productivity in each year (column (3)). We also control for field-times-year fixed effects which allow for cross-field (e.g., chemistry in 1915) differences in productivity in each year. Finally, we control for camp-times-field-times-year fixed effects which allow for cross-camp-and-field differences (e.g., chemistry in Allied countries in 1915) in productivity in each year. In the latter specification, the effect is identified from variation in the prewar reliance on foreign research between fields within the same camp, for example, biology in the United States, the United Kingdom, and other Allied

28. We measure top 1%, top 3%, and top 5% as the field-level percentiles within the 160 top scientific journals in our data. As these journals are the highest-cited journals of the time, the top 1% corresponds to an even more selected part of the overall citation distribution.

TABLE IX

EFFECT ON PUBLICATIONS: ROBUSTNESS CHECKS

Dependent variable: *Number of publications*	# pub. per author (1)	# pub. not normal. by # authors (2)	Control for camp × year (3)	Control for field × year (4)	Control for camp × field × year (5)
Panel A: Frontier measured by top 1%					
Prewar reliance on 1% frontier OUT × post-1914	−1.727***	−1.775**	−1.489*	−1.655***	−1.667***
	(0.638)	(0.669)	(0.826)	(0.616)	(0.491)
Prewar reliance on 1% frontier IN × post-1914	−0.827	−0.923	−0.823	−0.782	−0.762
	(0.736)	(0.730)	(0.764)	(0.725)	(0.698)
Within R-squared	0.062	0.066	0.064	0.064	0.068
Panel B: Frontier measured by top 3%					
Prewar reliance on 3% frontier OUT × post-1914	−0.784***	−0.813***	−0.596	−0.744**	−1.105***
	(0.282)	(0.288)	(0.379)	(0.295)	(0.237)
Prewar reliance on 3% frontier IN × post-1914	−0.363	−0.454	−0.432	−0.300	−0.311
	(0.283)	(0.279)	(0.297)	(0.292)	(0.265)
Within R-squared	0.062	0.066	0.063	0.064	0.068
Panel C: Frontier measured by top 5%					
Prewar reliance on 5% frontier OUT × post-1914	−0.380*	−0.400*	−0.224	−0.372	−0.686***
	(0.220)	(0.222)	(0.262)	(0.270)	(0.256)
Prewar reliance on 5% frontier IN × post-1914	−0.152	−0.205	−0.170	−0.120	−0.167
	(0.218)	(0.218)	(0.225)	(0.230)	(0.192)
Within R-squared	0.062	0.066	0.063	0.064	0.068
Scientist fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes			
Prewar reliance on nonfrontier	Yes	Yes	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes	Yes	Yes
Camp × year fixed effects			Yes		
Field × year fixed effects				Yes	
Camp × field × year fixed effects					Yes
Observations	227,084	227,084	227,084	227,084	227,084
Number of scientists	8,734	8,734	8,734	8,734	8,734

Notes. Each column and each panel reports one set of parameter estimates of regression (3) for the panel of university scientists. For the results presented in Panel A (B and C), we measure the frontier as the top 1% (top 3% and top 5%) research. The dependent variable measures the yearly number of publications in the 160 top journals in our data for the years 1905 to 1930. In column (1) and (3)–(5), the dependent variable is normalized by the number of authors and standardized to mean 0 and standard deviation 1 within fields. In column (2), the dependent variable is not normalized by the number of authors but standardized to mean 0 and standard deviation 1 within fields. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: $^{***}p < .01$, $^{**}p < .05$, and $^{*}p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

countries.[29] These fixed effects control for a differential impact of any shock that has the same effect on scientists' productivity in a certain field, camp, and year. For example, Allied chemists may have published less in 1919 because of their involvement in the organization of the boycott against Central scientists. Similarly, these fixed effects control for scientific breakthroughs that made scientists in a field and camp more productive in a certain year, for example, Central physicists (among them Werner Heisenberg from Germany and Erwin Schrödinger from Austria) at the height of the quantum revolution.

2. Potential Confounding Effects of WWI on Scientific Productivity. In additional specifications, we show that the results are presumably driven by a reduction in international knowledge flows, and not by more general disruption caused by WWI. While camp-times-field-times-year fixed effects control for yearly productivity shocks that affect all scientists in a certain camp and field, some war-related confounders may potentially be correlated with the dependence on frontier research from abroad.

While the war was not fought on the territories of the most important scientific powers (the United States, the United Kingdom, and Germany), it may nevertheless have disrupted scientific research in some field-country pairs because of other issues. For example, professors may have had fewer graduate students for joint projects because potential students were drafted, or professors may have been distracted by political engagement or by worries about the safety of their families and friends. Although we do not have direct measures for these confounders, we proxy for them with different country-level measures of war intensity that we interact with year fixed effects. The results are robust to controlling for an indicator of combat on the territory of the country, the total number of deaths per capita, the total number of civilian deaths per capita, and all three war-intensity measures at the same time (Table X, columns (1)–(4)).[30]

29. Differences in the prewar reliance on foreign research between fields within the same camp occur because field-country pairs produced different amounts of frontier research and because, even in normal times, frictions reduce knowledge flows across countries (see Online Appendix D for a stylized example of the identifying variation).

30. See Online Appendix E.5 for data sources on war-intensity measures.

TABLE X

EFFECT ON PUBLICATIONS: DISRUPTION OF KNOWLEDGE FLOWS OR OTHER WAR RELATED DISRUPTION?

Dependent variable: *Number of publications*	Control for combat in country (1)	Control for total deaths per capita (2)	Control for civilian deaths per capita (3)	Control for (1), (2), and (3) (4)	Exclude chemistry (5)	# pub. in own-camp journals (6)
Panel A: Frontier measured by top 1%						
Prewar reliance on 1% frontier OUT	−1.846***	−1.547***	−1.987***	−1.649***	−1.721***	−1.484**
× post-1914	(0.527)	(0.414)	(0.433)	(0.385)	(0.563)	(0.639)
Prewar reliance on 1% frontier IN	−0.739	−0.351	−0.489	−0.248	−0.653	−0.545
× post-1914	(0.731)	(0.742)	(0.690)	(0.666)	(0.740)	(0.707)
Within R-squared	0.068	0.068	0.068	0.069	0.069	0.068
Panel B: Frontier measured by top 3%						
Prewar reliance on 3% frontier OUT	−1.158***	−0.984***	−1.297***	−1.056***	−1.142***	−1.074***
× post-1914	(0.232)	(0.241)	(0.179)	(0.195)	(0.271)	(0.314)
Prewar reliance on 3% frontier IN	−0.297	0.024	−0.109	0.089	−0.281	−0.214
× post-1914	(0.272)	(0.312)	(0.268)	(0.278)	(0.281)	(0.301)
Within R-squared	0.068	0.068	0.068	0.069	0.069	0.068
Panel C: Frontier measured by top 5%						
Prewar reliance on 5% frontier OUT	−0.663**	−0.625**	−0.888***	−0.676**	−0.632**	−0.566*
× post-1914	(0.280)	(0.270)	(0.221)	(0.266)	(0.296)	(0.339)
Prewar reliance on 5% frontier IN	−0.176	−0.098	−0.061	−0.086	−0.060	−0.078
× post-1914	(0.191)	(0.205)	(0.177)	(0.189)	(0.202)	(0.212)
Within R-squared	0.068	0.068	0.068	0.068	0.069	0.068
Scientist fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Prewar reliance on nonfrontier	Yes	Yes	Yes	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes	Yes	Yes	Yes
Camp × field × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	227,084	227,084	227,084	227,084	197,782	215,046
Number of scientists	8,734	8,734	8,734	8,734	7,607	8,271

Notes. Each column and section reports one set of parameter estimates of regression (3) for the panel of university scientists. In columns (1)–(5), the dependent variable measures the yearly number of publications in all 160 top journals in our data in the years 1905 to 1930. In column (6), the dependent variable measures the yearly number of publications in own-camp journals in our data in the years 1905 to 1930. In all columns, the dependent variable is normalized by the number of authors and standardized to mean 0 and standard deviation 1 within fields. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: $^{***}p < .01$, $^{**}p < .05$, and $^{*}p < .1$. The data were collected by the authors and combine: scientist census data from *Minerva—Handbuch der Gelehrten Welt*, publication and citation data from *ISI Web of Science*, and war intensity data from Mouguel (2011), *1914–1918 online: International Encyclopedia of the First World War*, and Wikipedia (see Section III and Online Appendix E.5 for details).

To investigate whether the war had a more direct effect on the mortality of scientists in our sample, we collect data on more than 6,500 obituaries from *Science*, *Nature*, *Physikalische Zeitschrift*, *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, and *Kürschners Deutscher Gelehrten kalender* and match them to the sample of scientists with a university position by 1914 (see Online Appendix E.6 for details). In general, scientists in our sample did not die disproportionately during WWI (Online Appendix Figure A.12). Moreover, we also show that scientists reliant on frontier research from abroad did not die disproportionately during this period (Online Appendix Table A.14).

The results are also robust to excluding chemists, whose scientific productivity may have been differentially affected by research on chemical weapons during the war (Table X, column (5)). While camp-times-field-times-year fixed effects control for most of these changes, the propensity to engage in war-related research among chemists may have been correlated with the prewar reliance on frontier research from abroad. For example, U.K. chemists, who were more reliant on research from abroad, may have been more distracted by research on chemical weapons than U.S. chemists, who were less reliant on research from abroad.

Finally, scientists in field-country pairs that, in the prewar period, were heavily reliant on frontier research from outside the camp may have published more papers in journals from the other camp during normal times. Thus, in addition to contending with reduced international knowledge flows, these scientists may have faced greater difficulty in publishing their papers during a time of political hostility. We explore this possibility by focusing on publications in own-camp journals. The results remain unchanged (Table X, column (6)), presumably because the majority of scientists published in journals edited in their own camp (see Online Appendix Table A.4).

3. Field-level Variation within the United States. We also explore effects on productivity using variation across fields in the United States only. Some U.S. fields, such as biochemistry, relied on frontier research from outside the camp, while others, such as biology, relied mostly on frontier research from home. While the United States participated in WWI, no battles were fought on U.S. territory, and hence, war-related disruption may have impacted

U.S. scientists to a lesser extent. Furthermore, this analysis allows us to rule out that the results are driven by a general rise of U.S. science that may have been correlated with the prewar dependence on foreign research.

In the prewar period, the productivity of U.S. scientists in fields that relied on frontier research from outside the camp improved, relative to the productivity of scientists in fields that relied on frontier research from home (Figure X). After 1914, the productivity of scientists in fields that relied on frontier research from outside the camp declined sharply and did not recover until 1930. We test whether the trend-break in 1914 is statistically significant with a regression that includes linear trends and the interaction of each linear trend with a post-1914 indicator.[31] The estimated trend-break in 1914 for the "Prewar Reliance on Frontier OUT" has a p-value of .055. The productivity of U.S. scientists in fields that relied on frontier research from foreign countries inside the camp also improved in the prewar period. While the productivity of scientists in these fields continued to improve after 1914, it improved at a somewhat lower pace. The trend-break in 1914 for the "Prewar Reliance on Frontier IN" was smaller than for scientists in fields that relied on frontier research from outside the camp (p-value of .099).

V.B. Alternative Outcomes

The previous results indicate that scientists in field-country pairs that, in the prewar period, relied on frontier research from abroad, published significantly fewer papers in top journals after 1914. In the following section, we explore whether this decline in the quantity of research (published in top science journals) was associated with a decline in the impact of that research on basic science and technology. In this context, paper citations as a measure of impact are problematic because citations were heavily distorted during the war and the boycott, as highlighted in the

31. More specifically, we estimate regression (3) including linear trends for "Prewar Reliance on Frontier OUT," "Prewar Reliance on Frontier IN," and "Prewar Reliance on Frontier Home," plus nonfrontier trends and the interaction of each of these trends with a post-1914 indicator. We then test whether "Prewar Reliance on Frontier OUT" interacted with "post-1914" is significantly different from 0. The U.S. sample includes 11 fields and we cluster standard errors at the field level. To avoid a downward bias in estimated standard errors due to the small number of clusters (Cameron, Gelbach, and Miller 2008), we implement a cluster-bootstrap with asymptotic refinement as suggested by Cameron and Miller (2015).

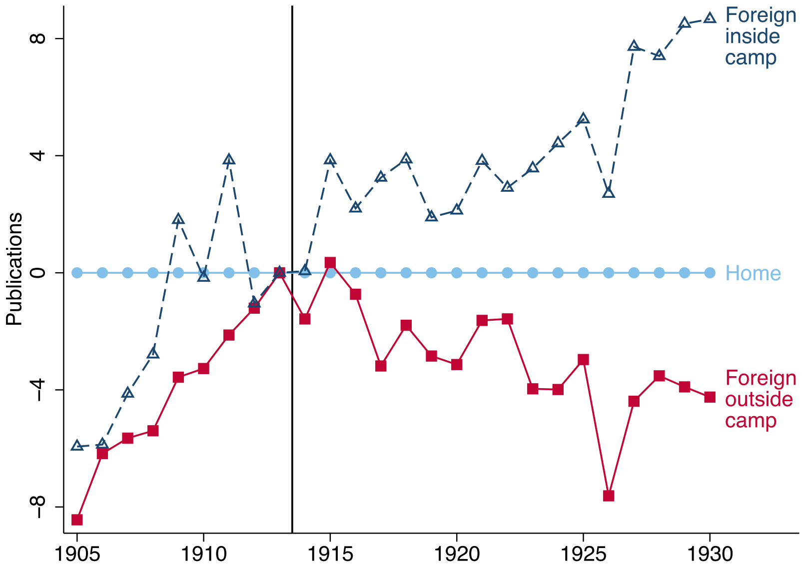

FIGURE X

Effect on Publications: Within U.S. Variation

The figure plots parameter estimates from regression (4) when we restrict the sample to scientists based in the United States. The "Foreign outside camp" line reports point estimates ($\beta_{1\tau}$) that measure changes in yearly publications for scientists in field-country pairs that, in the prewar period, relied on frontier research from outside the camp, compared to scientists who relied on frontier research from home. The "Foreign inside camp" line reports point estimates ($\beta_{2\tau}$) that measure changes in yearly publications for scientists in field-country pairs that, in the prewar period, relied on frontier research from foreign scientists inside the camp, compared to scientists who relied on frontier research from home. Prewar reliance on frontier research is measured by prewar citations to frontier research at the field-country pair level. Frontier research is defined as research that ended up in the top 1% of the subject-level citation distribution, counting citations until today. The regression also controls for prewar dependence on nonfrontier research from each camp interacted with year indicators. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt* and publication and citation data from *ISI Web of Science* (see Section III for details).

first part of the paper. We therefore investigate effects on three new measures of research impact: Nobel-nominated research, scientific research that introduced novel words, and a measure of how often these words were applied in U.S. patents.

1. Nobel-Nominated Research. To investigate effects on path-breaking research, we analyze changes in the probability of producing research that led to a Nobel Prize nomination. The Nobel Prize has been awarded by the Academy of Sciences and the Karolinska Institutet in Sweden, a Neutral country.

We collect data on all nominations for the physics, chemistry, and physiology/medicine prizes from the Nobel Nomination Archive (see Nobelprize.org 2014). Between 1905 and 1945, 993 individuals were nominated for a Nobel Prize at least once, and 131 of them eventually won it. The database does not list the exact research that led to a nomination. We identify that research by searching our publication data for the highest-cited paper (counting citations until today) that a nominee published before his last nomination (see Online Appendix E.4 for details).[32] We then generate an indicator, "Nobel-nominated paper" that equals 1 if a scientist published his "Nobel-nominated paper" in a certain year, and 0 for all other years.

For example, Arthur Compton received the 1927 Nobel Prize in physics "for the discovery of the effect named after him." He was last nominated for the prize in 1927, and we therefore search for the highest-cited paper he published before 1927. His article, "A Quantum Theory of the Scattering of X-Rays by Light Elements," was published in the *Physical Review* in 1923, and received (until today) 355 citations, more than any other of his pre-1927 papers. For Arthur Compton the "Nobel-nominated paper" indicator therefore equals 1 in 1923 and 0 in all other years.

While some candidates "only" received one nomination for the Nobel Prize, others received many more. To distinguish papers at the very highest level of the quality spectrum, we construct a second measure that weights the Nobel-nominated papers by the number of nominations. Because scientists who eventually won

32. Ideally, we would not rely on citations to identify the year of Nobel-nominated research. However, this year cannot be systematically identified from the data posted by the Nobel archives. Jones and Weinberg (2011) collect biographical data to identify the period of key research for Nobel Prize winners. Our measure of Nobel-nominated research identifies a single year. For Nobel Prize winners, our measure has a correlation of 0.69 with the middle year of the period of key research reported by Jones and Weinberg. The detailed information that Jones and Weinberg use to construct their measure is not available for scientists who were nominated but did not win.

the prize experienced a hike in nominations in the last two years before winning (Online Appendix Figure A.14), we focus on the number of nominations during the last two years before a candidate's last nomination. The physicists with the highest number of nominations in the last two years were Albert Einstein (31 nominations), Jean Perrin (18), Werner Heisenberg (17), and Erwin Schrödinger (17); they all eventually won a Nobel Prize, and they are considered to have made some of the most outstanding contributions to physics in this period. The measure is highly predictive of winning the Nobel Prize.[33] Candidates with one nomination only had a 4% chance of winning. Candidates with two nominations had a 13% chance, candidates with three nominations had a 16% chance, candidates with four nominations had a 19% chance, candidates with five to nine nominations had a 40% chance, and candidates with more than nine nominations had a 61% chance of winning (Online Appendix Figure A.13).

Using the "Nobel-nominated paper" variable as the dependent variable, we estimate regression (3) for our sample of university scientists.[34] After 1914, the probability of publishing a Nobel-nominated paper declined significantly for scientists in field-country pairs that relied on frontier research (measured by the top 1%) from outside the camp (Table XI, Panel A, column (1), significant at the 5% level). The estimated effect indicates that the probability of publishing a Nobel-nominated paper declined by 0.001 for scientists in a field such as U.S. biochemistry that relied heavily on frontier research from outside the camp, compared to scientists in a field such as U.S. biology that relied mostly on frontier research from home. The prewar period probability of writing a Nobel-nominated paper in fields that relied on frontier research from abroad is also 0.001. Thus, the results indicate that the decline in international scientific cooperation effectively wiped out the chance of writing a paper worthy of a Nobel Prize nomination for scientists in field-country pairs reliant on frontier research from outside the camp. The results are robust to using

33. The number of nominations in the last two years before the last nomination is a better predictor of winning than the total number of nominations, because the total number of nominations is censored for winners (i.e., most of them were no longer nominated after winning).

34. The estimation includes 234 nominees, among them 42 winners. Of the 993 potential nominees, 474 published their Nobel-nominated paper between 1905 and 1930, and 234 of them had a university position by 1914.

TABLE XI

EFFECT ON NOBEL-NOMINATED PAPERS

Dependent variable:	Nomination paper		Nomination paper weighted by # noms.	
	(1)	(2)	(3)	(4)
Panel A: Frontier measured by top 1%				
Prewar reliance on 1% frontier OUT	−0.021**	−0.019*	−0.148***	−0.175*
× post-1914	(0.008)	(0.011)	(0.052)	(0.103)
Prewar reliance on 1% frontier IN	−0.005	−0.003	−0.048	−0.033
× post-1914	(0.008)	(0.009)	(0.041)	(0.062)
Within R-squared	0.001	0.005	0.001	0.004
Panel B: Frontier measured by top 3%				
Prewar reliance on 3% frontier OUT	−0.012***	−0.012***	−0.061**	−0.073
× post-1914	(0.004)	(0.004)	(0.028)	(0.045)
Prewar reliance on 3% frontier IN	−0.005	−0.006	−0.021	−0.011
× post-1914	(0.004)	(0.005)	(0.018)	(0.020)
Within R-squared	0.001	0.005	0.001	0.004
Panel C: Frontier measured by top 5%				
Prewar reliance on 5% frontier OUT	−0.010**	−0.010**	−0.072**	−0.074**
× post-1914	(0.004)	(0.004)	(0.027)	(0.030)
Prewar reliance on 5% frontier IN	−0.002	−0.003	0.012	0.020
× post-1914	(0.003)	(0.003)	(0.017)	(0.019)
Within R-squared	0.001	0.005	0.001	0.004
Scientist fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes		Yes	
Prewar reliance on nonfrontier	Yes	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes	Yes
Camp × field × year fixed effects		Yes		Yes
Observations	227,084	227,084	227,084	227,084
Number of scientists	8,734	8,734	8,734	8,734

Notes. Each column and each panel reports one set of parameter estimates of regression (3) for the panel of university scientists. The dependent variable in columns (1) and (2) is an indicator that equals 1 if a scientist published a Nobel-nominated paper in a certain year between 1905 and 1930, and 0 for all other years. The dependent variable in columns (3) and (4) weights the Nobel-nominated paper indicator by the number of nominations in the two years before a candidate's last nomination. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: *** $p < .01$, ** $p < .05$, and * $p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt*, publication and citation data from *ISI Web of Science*, and Nobel nomination data from Nobelprize.org (2014) (see Section III for details).

different definitions of frontier research (Panels A to C) and to adding camp-times-field-times-year fixed effects (column (2)). We obtain similar results if we weight the Nobel-nominated indicator by the number of nominations (Table XI, columns (3) and (4)).

These results suggest that access to the very best research, especially the top 1%, is key for the production of path-breaking ideas.

2. Novel Scientific Words. As an alternative outcome, we count the number of novel words that a scientist introduced to the scientific community in each year. The measure proxies for the introduction of new scientific concepts that required new scientific terms. We define novel words as words that the scientist first used in a title of a paper published between 1905 and 1930, and that had not been used in any prior paper title. To check whether a word had been used before, we not only consider the papers published by the scientists in our estimation sample, but all papers that were published in any of the 160 journals in the *Web of Science* between 1900 and 1930. To ensure that we do not consider words that were already commonly used in other domains, we exclude frequently used words, as well as all numbers, from the data.[35] As above, we standardize the outcome variable to have mean zero and unit variance within fields.

One example of a novel word is "magnetron," which was introduced by U.S. physicist Albert W. Hull in the paper "The Measurement of Magnetic Fields of Medium Strength by Means of a Magnetron," published in the *Physical Review* in 1923. Another example is "electroencephalogram," which was introduced by German psychiatrist Hans Berger in the paper "Electroencephalogram of Humans," published in the *Journal für Psychologie und*

35. We exclude the 10,000 most frequently used words in English-language books contained in the Project Gutenberg database as of April, 16 2006 (available at https://en.wiktionary.org/wiki/Wiktionary:Frequency_lists#English). Project Gutenberg currently contains the full text of over 53,000 books. Because the database contains books whose copyright have expired, the typical book in the database was published before 1923. The most frequently used words therefore reflect historical language use that is more relevant for the period of our analysis. The results are robust to excluding only 5,000 or all 36,662 frequently used words (see Online Appendix Table A.15). For the main results, we do not remove all frequently used words because words such as quantum (on position 17,132) may have existed before but might have taken on a new meaning with the publication of a scientific paper. For more detail on the novel scientific words measure see Online Appendix E.4.

Neurologie in 1930.[36] Other examples of novel words that were introduced in this period are hormone, isotope, superconductor, and chemical substances such as 5-trinitro-4-acetylaminophenol. Introducing novel words is rare; the average scientist introduced 0.042 novel words per year (Table VII).[37]

Using the number of novel words as the dependent variable, we estimate regression (3). After 1914, scientists in field-country pairs that relied on frontier research (measured by the top 1%) from outside the camp published fewer papers that introduced novel words (Table XII, Panel A, column (1), significant at the 1% level).[38] The estimated effect indicates that scientists in a field such as U.S. biochemistry, that relied heavily on frontier research from outside the camp, introduced 0.07 standard deviations fewer words than scientists in a field such as U.S. biology that relied mostly on frontier research from home. Scientists who relied on frontier research from inside the camp also published fewer papers that introduced novel words (Table XII, Panel A, column (1), significant at the 10% level). When we measure the frontier with the top 3% or top 5% of research, we only find significant effects in specifications that add camp-times-field-times-year fixed effects (Table XII, Panels B and C, columns (1) and (2)). As before, the results are strongest if we measure the frontier with the top 1% of research, suggesting that access to the top 1% is particularly important to produce papers that introduce new scientific concepts.

3. Novel Scientific Words that Are Applied in Technology. We also investigate an outcome that measures how basic science was

36. Scientists typically publish a number of papers summarizing an important discovery. We count the first appearance of a novel word in the 160 journals in our data, which may not necessarily be the very first time the word appeared in any scientific publication. Albert Hull, for example, published "The Magnetron" in the *Journal of the American Institute of Electrical Engineers* in September 1921. This journal is not in our data because it was not a core journal for scientists in mathematics, physics, chemistry, biology, or medicine.

37. The total number of novel scientific words that a scientist introduced in papers published between 1905 and 1913 has a correlation of 0.51 with the total number of citations that these papers have received until today.

38. The number of novel words would increase artificially if scientists in opposing camps started to use different terms for the same scientific concept after 1914. Our measure of novel words is less susceptible to this concern because ISI translated all paper titles into English in 2004. Furthermore, more international field-country pairs should be more exposed to such an artificial increase in the number of words. This would bias our estimates toward zero.

TABLE XII

EFFECT ON WORD INNOVATION AND PATENTS

Dependent variable:	Novel scientific words		Patent-relevant words	
	(1)	(2)	(3)	(4)
Panel A: Frontier measured by top 1%				
Prewar reliance on 1% frontier OUT	−1.229***	−0.778**	−1.134***	−1.207***
× post-1914	(0.441)	(0.327)	(0.349)	(0.297)
Prewar reliance on 1% frontier IN	−0.910*	−0.918**	−0.569*	−0.661**
× post-1914	(0.468)	(0.368)	(0.295)	(0.253)
Within R-squared	0.025	0.028	0.015	0.018
Panel B: Frontier measured by top 3%				
Prewar reliance on 3% frontier OUT	−0.311	−0.359**	−0.415**	−0.639***
× post-1914	(0.261)	(0.173)	(0.198)	(0.155)
Prewar reliance on 3% frontier IN	−0.149	−0.181	−0.115	−0.129
× post-1914	(0.225)	(0.167)	(0.182)	(0.139)
Within R-squared	0.024	0.028	0.015	0.018
Panel C: Frontier measured by top 5%				
Prewar reliance on 5% frontier OUT	−0.182	−0.298*	−0.339**	−0.542***
× post-1914	(0.204)	(0.164)	(0.158)	(0.149)
Prewar reliance on 5% frontier IN	−0.149	−0.150	−0.136	−0.161
× post-1914	(0.173)	(0.134)	(0.134)	(0.102)
Within R-squared	0.024	0.028	0.015	0.018
Scientist fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes		Yes	
Prewar reliance on nonfrontier	Yes	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes	Yes
Camp × field × year fixed effects		Yes		Yes
Observations	227,084	227,084	227,084	227,084
Number of scientists	8,734	8,734	8,734	8,734

Notes. Each column and each panel reports one set of parameter estimates of regression (3) for the panel of university scientists. The dependent variable in columns (1) and (2) counts the number of novel words that appeared in the title of a scientific paper published in year t. The dependent variable in columns (3) and (4) counts the number of times each of the novel words (as defined above) was used in the text of any patent granted by the U.S. Patent Office in years $t + 15$ and $t + 30$. The dependent variable in columns (3) and (4) is winsorized at the 99th percentile. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: *** $p < .01$, ** $p < .05$, and * $p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt*, publication and citation data from *ISI Web of Science*, Nobel nomination and award data from Nobelprize.org (2014), and patent data from the *U.S. Patent Office* (see Section III for details).

applied in the development of new technologies. Specifically, we measure how often subsequent patents used the novel words that were introduced to the scientific community (as described in the previous section). We obtain the full text of 2.5 million patents that were granted between 1920 and 1979 from the U.S. Patent Office web page.[39] We then search the 7.5 billion words in these patents for the novel words that scientists in our sample introduced to the scientific community. For a paper published in year t, the measure counts the number of times a novel scientific word appears in subsequent patents that were granted between year $t + 15$ and $t + 30$. As an example, for a paper published in year $t = 1905$ that introduced a novel word, we search patents granted between 1920 and 1935.[40] This measure of patent-relevant words weights the novel scientific words introduced by each scientist in a certain year with the number of times these words appeared in subsequent patent grants.[41]

For example, the novel scientific word "magnetron," which was introduced in 1923, appeared 9,538 times in 997 patents after 1923. The magnetron was later used to dramatically improve radar technology. It serves at the heart of microwave ovens, and provides the key underlying technology for sulfur lamps. Examples of patents that use the word "magnetron" are U.S. patent no. 2,115,521 "Magnetron Oscillator and Detector" granted in 1939, and U.S. patent no. 2,605,383 "Means of treating foodstuffs," one

39. Ideally, we would search patents that were granted from 1905 onward, but unfortunately the text of 1905 to 1920 patents is not available in digitized form.

40. The time window ensures that we measure the link between basic science and technology with a consistent time delay. This accounts for the fact that patent data are only available from 1920 onwards. As a result of this data limitation, novel words that were introduced in scientific papers published in 1905 can only be observed in patents after 15 years. While the data structure leads us to measure effects of basic science on patenting with a 15-year delay, earlier research shows that basic pharmaceutical research is associated with U.S. Food and Drug Administration approval of new molecular entities with a delay of 17 to 24 years (Toole 2012), and that the stock of basic science affects total factor productivity growth with a delay of about 20 years (Adams 1990). The results are robust to considering all patents granted between 1920 (or the publication year of the relevant paper if it was published after 1920) and 1979 (see Online Appendix Table A.16, columns (11) and (12)).

41. This measure may overstate the effect of basic science on the development of new technology if certain novel words appeared independently of each other in papers and patents. As long as independent discoveries did not change differentially across field-country pairs over time, the estimates of the effect of reduced international cooperation would remain unbiased.

of the first microwave patents granted in 1952. The novel word "electroencephalogram," which was introduced in 1930, appeared seven times in three patents after 1930. The electroencephalogram allows monitoring of electrical activity in the brain, and it is used to diagnose epilepsy, coma, and brain death. In the past, it was used to diagnose tumors and stroke, but this use declined with the invention of computed tomography and magnetic resonance imaging scans. An example of a patent that uses the word "electroencephalogram" is U.S. patent no. 2,409,033 "Electroencephalograph device," granted in 1946.

Per year, the average scientist introduced novel words that appeared 0.43 times in subsequent patents. The number of novel scientific words that are applied in new technology is highly skewed because most scientists never introduced a novel scientific word, but a few scientists introduced words that were frequently applied in patents, for example "magnetron" appeared 9,538 times. To avoid results driven by a few outliers we winsorize the outcome variable at the 99th percentile.[42] We also standardize the outcome variable to have mean zero and unit variance within fields.

Using patent-relevant words as the dependent variable, we estimate regression (3). After 1914, scientists in field-country pairs that relied on frontier (measured by the top 1%) research from outside the camp published less scientific research that introduced novel words relevant for patenting (Table XII, Panel A, column (3), significant at the 1% level). The estimated effect indicates that scientists in a field such as U.S. biochemistry, that relied heavily on frontier research from outside the camp, reduced patent-relevant words by 0.05 standard deviations compared to scientists in a field such as U.S. biology that relied mostly on frontier research from home. Scientists reliant on frontier research from inside the camp also reduced patent-relevant words, although this result is only significant if we control for camp-times-field-times-year fixed effects (Table XII, Panel A, columns (3) and (4)). When we measure the frontier with the top 3% or top 5% of research, we only find significant reductions in patent-relevant words for scientists in fields that predominantly relied on frontier research from outside the camp (Table XII, Panels B and C, columns (3) and (4)).

42. The results are similar if we do not winsorize the data (see Online Appendix Table A.16).

Patent-relevant words may have declined after 1914 in field-country pairs reliant on frontier research from abroad, for two main reasons. First, scientists in field-country pairs reliant on frontier research from abroad introduced fewer novel scientific words. As a consequence, the papers of these scientists could have been less useful for inventors. As a result, the scientists' patent-relevant word measure would have decreased. Second, inventors themselves may have lost access to basic science from abroad. If inventors had sourced basic science research similarly to scientists, the patent-relevant word measure would have increased for scientists in field-country pairs reliant on frontier research from abroad.[43] As we find a relative decline in patent-relevant words for scientists in field-country pairs reliant on frontier research from abroad, the results indicate that these scientists introduced fewer novel words that were useful for inventors. This suggests that the decline in international scientific cooperation not only affected the production of basic science but also impeded the application of basic science in the development of new technologies.

V.C. Who Benefits from Access to Frontier Research

Finally, we investigate whether high- or low-quality scientists are differentially affected by the reduction in international scientific cooperation. We split the sample into high- and low-quality scientists according to the field-level median productivity in the prewar period, as measured by publications in the 160 top journals in our data. We then separately estimate regression (3) for high- and low-quality scientists.

Output of above-median productivity scientists decreased 5 to 15 times more, in absolute terms, than output of below-median scientists (Table XIII). These findings are consistent across all outcomes (columns (1)–(10)) and the different definitions of frontier research (Panels A to C). We measure output in absolute terms in

43. The following example illustrates the second channel: consider U.S. inventors who lost access to basic science from Germany. If U.S. biochemistry inventors had relied on basic science from Germany and U.S. biology inventors had relied on basic science from home, U.S. biochemistry inventors would have disproportionately reduced the application of basic science from Germany. As a result, German biochemistry scientists would have experienced a disproportionate decline in the application of their novel scientific words by U.S. inventors. Hence, German biology scientists would have experienced an increase in their patent-relevant word measure, relative to German biochemistry scientists.

TABLE XIII

EFFECT ON PRODUCTIVITY BY PREWAR QUALITY OF SCIENTISTS

	⩽ median (1)	> median (2)	⩽ median (3)	> median (4)	⩽ median (5)	> median (6)	⩽ median (7)	> median (8)	⩽ median (9)	> median (10)
Dependent variable:	Number of publications		Nomination paper		Nom. paper (weighted)		Novel scientific words		Patent-relevant words	
Panel A: Frontier measured by top 1%										
Prewar reliance on 1% frontier OUT	−0.430*	−3.295***	0.004	−0.052**	−0.023	−0.373*	−0.149	−1.656**	−0.465***	−2.163***
× post-1914	(0.227)	(1.237)	(0.007)	(0.022)	(0.039)	(0.219)	(0.099)	(0.778)	(0.135)	(0.714)
Prewar reliance on 1% frontier IN	0.028	−1.693	−0.000	−0.010	−0.004	−0.083	0.026	−1.884***	−0.134	−1.148*
× post-1914	(0.500)	(1.046)	(0.006)	(0.017)	(0.030)	(0.125)	(0.167)	(0.528)	(0.187)	(0.587)
Within R-squared	0.124	0.077	0.008	0.009	0.016	0.006	0.025	0.046	0.014	0.031
Panel B: Frontier measured by top 3%										
Prewar reliance on 3% frontier OUT	−0.278**	−2.131***	0.001	−0.031***	−0.004	−0.166	−0.164***	−0.540	−0.327***	−0.992***
× post-1914	(0.120)	(0.569)	(0.002)	(0.011)	(0.013)	(0.100)	(0.056)	(0.347)	(0.072)	(0.300)
Prewar reliance on 3% frontier IN	−0.088	−0.668	−0.001	−0.012	−0.004	−0.016	−0.037	−0.421	−0.136	−0.125
× post-1914	(0.182)	(0.438)	(0.003)	(0.008)	(0.013)	(0.034)	(0.073)	(0.294)	(0.096)	(0.290)
Within R-squared	0.124	0.077	0.008	0.009	0.016	0.006	0.025	0.046	0.014	0.031
Panel C: Frontier measured by top 5%										
Prewar reliance on 5% frontier OUT	−0.070	−1.286***	−0.001	−0.024**	−0.023	−0.142**	−0.140**	−0.295	−0.290***	−0.714***
× post-1914	(0.167)	(0.454)	(0.003)	(0.010)	(0.015)	(0.065)	(0.067)	(0.306)	(0.079)	(0.266)
Prewar reliance on 5% frontier IN	−0.057	−0.392	−0.000	−0.008	0.001	0.036	−0.057	−0.335	−0.151**	−0.214
× post-1914	(0.150)	(0.293)	(0.002)	(0.005)	(0.010)	(0.036)	(0.055)	(0.214)	(0.064)	(0.178)
Within R-squared	0.124	0.077	0.008	0.009	0.016	0.006	0.025	0.046	0.014	0.031
Scientist fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prewar reliance on nonfrontier	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Career age × field interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Camp × field × year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	121,446	105,638	121,446	105,638	121,446	105,638	121,446	105,638	121,446	105,638
Number of scientists	4,671	4,063	4,671	4,063	4,671	4,063	4,671	4,063	4,671	4,063

Notes. Each column and each panel reports one set of parameter estimates of regression (3) for the panel of university scientists. We divide the sample of scientists into two subsamples: scientists who published less or the same number of papers than the median scientist (results reported in odd columns) and scientists who published more papers than the median scientist (even columns). The median number of papers is calculated at the field level for the period 1905–1913. The dependent variables are defined as in the previous tables. "Prewar reliance on frontier OUT" is the prewar citation share to frontier research (1%, 3%, or 5%) from outside the camp. "Prewar reliance on frontier IN" is the prewar citation share to frontier research (1%, 3%, or 5%) from foreign countries inside the camp. The reference/omitted category is "Prewar reliance on frontier HOME." Standard errors are clustered at the country-times-field level. Significance levels: *** $p < .01$, ** $p < .05$, and * $p < .1$. The data were collected by the authors and combine scientist census data from *Minerva—Handbuch der Gelehrten Welt*, publication and citation data from *ISI Web of Science*, Nobel nomination data from Nobelprize.org, and patent data from the U.S. Patent Office (see Section III for details).

order to capture overall scientific progress. However, relative to prewar means, output declined relatively more for below-median scientists.[44]

These results suggest a complementarity between access to frontier research and the underlying quality of scientists.

VI. Conclusion

The dramatic decline in international scientific cooperation around WWI enables us to study how frontier research affects scientific productivity. This sheds light on the importance of path-breaking research circulating among the most fertile minds in academic communities worldwide. Because our results suggest that access to frontier research is key for the production of ideas, including path-breaking ones, one can conclude that facilitating access to frontier research can substantially increase the production of basic science. Access needs to be interpreted in a broad sense: not only physical access to journal articles, conferences, and research seminars, but also discerning the thin, ever-advancing, and truly path-breaking edge of the frontier from the millions of scientific papers published every year.

Our results suggest that science policy should therefore be geared toward facilitating access to and capitalizing on the potential catalytic effects of frontier research in enhancing scientific progress. Providing open access to journals may partly achieve this goal. However, discerning what constitutes frontier research requires skills that are hard to develop without guidance from leading scientists working at the forefront of scientific endeavor. Personal contacts are particularly useful because face-to-face interactions are a superior way of transmitting ideas (e.g., Glaeser 2011; Head, Li, and Minondo 2015). High-quality PhD programs at universities where frontier research proliferates can therefore help put young scientists on the most promising career paths (Waldinger 2010). Even more established scientists can profit from long and short-term visits at the centers of science (Catalini, Fons-Rosen, and Gaule 2016) and from attending high-quality conferences (de Leon and McQuillin 2015) and research

44. Due to the very low prewar means of below-median scientists, we find larger relative changes for these scientists. If, alternatively, we measure changes relative to 1905–1930 means, relative changes are larger for above-median scientists.

seminars. A famous example of a fruitful interaction between researchers is the series of lectures that Danish physicist and Nobel laureate Niels Bohr held at Göttingen in 1922, sometimes dubbed the "Bohr Festival." At this event, Bohr presented his latest theories of atomic structure and exchanged ideas with his peers, including (future) Nobel laureates James Franck, Max Born, Wolfgang Pauli, and the young physics prodigy Werner Heisenberg (e.g., Mehra and Rechenberg 1982, 345). In fact, Bohr underscored that being from a small country made it even more important to interact with international scientists producing frontier research (Bohr 2007, 172).

Our results also suggest that access to frontier research not only affects the production of basic science; it also increases the application of science in the development of new technology. Hence, policies that widen access to frontier research could benefit society beyond the confines of science itself.

University of Bristol
University of Warwick
London School of Economics

Supplementary Material

An Online Appendix for this article can be found at The Quarterly Journal of Economics online. Data and code replicating tables and figures in this article can be found in Iaria, Schwarz, and Waldinger (2017), in the Harvard Dataverse, doi:10.7910/DVN/SH1KE7.

References

Abramitzky, R., and I. Sin, "Book Translations as Idea Flows: The Effects of the Collapse of Communism on the Diffusion of Knowledge," *Journal of the European Economic Association*, 12 (2014), 1453–1520.

Acs, Zoltan J., David B. Audretsch, and Maryann P. Feldman, "Real Effects of Academic Research: Comment," *American Economic Review*, 82 (1992), 363–367.

Adams, James D., "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy*, 98 (1990), 673–702.

Aghion, P., L. Boustan, C. Hoxby, and J. Vandenbussche, "The Causal Impact of Education on Economic Growth: Evidence from U.S." (2009).

Agrawal, Ajay K, John McHale, and Alex Oettl, "How Stars Matter: Recruiting and Peer Effects in Evolutionary Biology," *Research Policy*, 46 (2017), 853–867.

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

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat, “Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules,” Mimeo, 2016.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, “Superstar Extinction,” *Quarterly Journal of Economics*, 125 (2010), 549–589.

Belenzon, Sharon, and Mark Schankerman, “Spreading the Word: Geography, Policy, and Knowledge Spillovers,” *Review of Economics and Statistics*, 95 (2013), 884–903.

Biasi, B., and P. Moser, “The Impact of Access on Science Evidence from the WWII Book Republication Program,” Working paper, New York University, 2015.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan, “Latent Dirichlet Allocation,” *Journal of Machine Learning Research*, 3 (2003), 993–1022.

Bohr, Niels, *Collected Works: Popularization and People (1911–1962)*, vol. 12. (Dordrecht: Elsevier 2007).

Borjas, George J., and Kirk B. Doran, “The Collapse of the Soviet Union and the Productivity of American Mathematicians,” *Quarterly Journal of Economics*, 3 (2012), 1143–1203.

———, “Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors,” *Review of Economics and Statistics*, 97 (2015), 1104–1117.

Bryan, Kevin A., and Yasin Ozcan, “The Impact of Open Access Mandates on Invention,” Mimeo, Toronto, 2016.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 90 (2008), 414–427.

Cameron, A. Colin, and Douglas L. Miller, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 50 (2015), 317–372.

Catalini, Christian, Christian Fons-Rosen, and Patrick Gaule, “Did Cheaper Flights Change the Direction of Science?” IZA Discussion Paper no. 9897 2016.

Cock, A. J., “Chauvinism and Internationalism in Science: The International Research Council, 1919–1926,” *Notes and Records of the Royal Society of London*, 37 (1983), 249–288.

Crawford, E., *Nationalism and Internationalism in Science, 1880–1939: Four Studies of the Nobel Population* (Cambridge: Cambridge University Press, 1988).

Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman, “Indexing by Latent Semantic Analysis,” *Journal of the American Society for Information Science*, 41 (1990), 391.

de Leon, Fernanda, and Ben McQuillin, “The Role of Conferences on the Pathway to Academic Impact: Evidence from a Natural Experiment,” Mimeo, University of Kent, 2015.

Forschungen und Fortschritte, “Deutsche Wissenschaft und Ausland in der Statistik. I. Wissenschaftliche Kongresse und Organisationen,” *Forschungen und Fortschritte*, 9 (1933), 330–332.

Furman, Jeffrey L., and Scott Stern, “Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research,” *American Economic Review*, 101 (2011), 1933–1963.

Galasso, A., and M. Schankerman, “Patents and Cumulative Innovation: Causal Evidence from the Courts,” *Quarterly Journal of Economics*, 130 (2015), 317–369.

Glaeser, Edward, *Triumph of the City: How Urban Spaces Make Us Human* (London: Pan Macmillan 2011).

Greenaway, Frank, *Science International: A History of the International Council of Scientific Unions* (Cambridge: Cambridge University Press 1996).

Head, Keith, Yao Amber Li, and Asier Minondo, “Geography, Ties, and Knowledge Flows: Evidence from Citations in Mathematics,” available at SSRN 2660041, (2015).

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger, “Replication Data for: ‘Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science’,” 2017. doi:10.7910/DVN/SH1KE7.

Jaffe, A. B., M. Trajtenberg, and R. Henderson, “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *Quarterly Journal of Economics*, 108 (1993), 577–598.

Jaffe, Adam B., “Real Effects of Academic Research,” *American Economic Review*, (1989), 957–970.

Jones, B. F., “The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?,” *Review of Economic Studies*, 76 (2009), 283–317.

Jones, B. F., and B. A. Weinberg, “Age Dynamics in Scientific Creativity,” *Proceedings of the National Academy of Sciences*, 108 (2011), 18910–18914.

Jones, Charles I., “R & D-based Models of Economic Growth,” *Journal of Political Economy*, 103 (1995), 759–784.

Kerkhof, K., “Das Versailler Diktat und die Deutsche Wissenschaft: Ein Beitrag zur Geschichte der Internationalen Organisationen,” *Monatshefte für Auswärtige Politik*, 7 (1940), 836–850.

Kevles, D. J., “‘Into Hostile Political Camps’: The Reorganization of International Science in World War I,” *Isis*, 62 (1971), 47–60.

Landauer, Thomas K., Peter W. Foltz, and Darrell Laham, “An Introduction to Latent Semantic Analysis,” *Discourse Processes*, 25 (1998), 259–284.

Lee, Daniel D., and H. Sebastian Seung, “Algorithms for Non-negative Matrix Factorization,” in *Advances in Neural Information Processing Systems*, (IEEE Conference on Neural Information Processing Systems: Natural and Synthetic, Massachusetts Institute of Technology Press, 2001) 556–562.

Lehto, O., *Mathematics without Borders: A History of the International Mathematical Union* (Berlin: Springer Science and Business Media 1998).

Mansfield, Edwin, “Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing,” *Review of Economics and Statistics*, 77 (1995), 55–65.

Martin, Dian I., and Michael W. Berry, “Mathematical Foundations behind Latent Semantic Analysis,” In *Handbook of Latent Semantic Analysis*, edited by T. K. Landauer. Mahwah, NJ: Lawrence Erlbaum Associates, 2017, 35–56.

Mehra, Jagdish, *The Solvay Conference on Physics: Aspects of the Development of Physics since 1911* (Dordrecht: Reidel 1975).

Mehra, Jagish, and Helmut Rechenberg, *The Historical Development of Quantum Theory*, vol. 1. (Dordrecht: Spring Science and Business Media 1982).

Mokyr, Joel, *The Gifts of Athena: Historical Origins of the Knowledge Economy* (Princeton, NJ: Princeton University Press 2002).

Moser, P., A. Voena, and F. Waldinger, “German Jewish Émigrés and U.S. Invention,” *American Economic Review*, 104 (2014), 3222–3255.

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

Mougel, Nadege, *REPERES: World War I Casualties* (Scy-Chazelles, France: CERS, 2011).

Mukherjee, Satyam, Daniel M. Romero, Ben Jones, and Brian Uzzi, “The Nearly Universal Link between the Age of Past Knowledge and Tomorrow’s Breakthroughs in Science and Technology: The Hotspot,” *Science Advances*, 3 (2017), e1601315.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, “The Allocation of Talent: Implications for Growth,” *Quarterly Journal of Economics*, 106 (1991), 503–530.

Murray, F., P. Aghion, M. Dewatripont, J. Kolev, and S. Stern, “Of Mice and Academics: Examining the Effect of Openness on Innovation,” NBER Working Paper no. 14819, 2009.

Newton, Isaac, “Letter from Isaac Newton to Robert Hooke,” (1675), http://digitallibrary.hsp.org/index.php/Detail/Object/Show/object_id/9285.

Nobelprize.org, “Nomination Archive. Explore the Nomination Databases in Physics, Chemistry, Physiology or Medicine, Literature and Peace,” (2014), http://www.nobelprize.org/nomination/archive/.

Oettl, Alexander, “Reconceptualizing Stars: Scientist Helpfulness and Peer Performance,” *Management Science*, 58 (2012), 1122–1140.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, 12 (2011), 2825–2830.

Peri, Giovanni, “Determinants of Knowledge Flows and Their Effect on Innovation,” *Review of Economics and Statistics*, 87 (2005), 308–322.

Porter, Martin F., “An Algorithm for Suffix Stripping,” *Program*, 14 (1980), 130–137.

———, “Snowball: A Language for Stemming Algorithms,” (2001), http://snowball.tartarus.org/texts/.

Ramsay, W., “Germany’s Aims and Ambitions,” *Nature*, 94 (1914), 137–139.

Reid, C., *David Hilbert* (New York: Springer 1970).

Reinbothe, R., *Deutsch als Internationale Wissenschaftssprache und der Boykott nach dem Ersten Weltkrieg* (Peter Lang, Frankfurt am Main, 2006).

Romer, P. M., “Endogenous Technological Change,” *Journal of Political Economy*, 98 (1990), S71–S102.

Romer, Paul M., “Increasing Returns and Long-Run Growth,” *Journal of Political Economy*, 94 (1986), 1002–1037.

Schroeder-Gudehus, B., “Challenge to Transnational Loyalties: International Scientific Organizations after the First World War,” *Science Studies*, 3 (1973), 93–118.

Scotchmer, Suzanne, “Standing on the Shoulders of Giants: Cumulative Research and the Patent Law,” *Journal of Economic Perspectives*, 5 (1991), 29–41.

Stevens, Keith, Philip Kegelmeyer, David Andrzejewski, and David Buttler, “Exploring Topic Coherence over Many Models and Many Topics,” in *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (Stroudsburg, PA, USA: Association for Computational Linguistics, 2012), 952–961.

The Economist, “Who Owns the Knowledge Economy?” (*The Economist*, April 8th, 2000).

Thompson, P., and M. Fox-Kean, “Patent Citations and the Geography of Knowledge Spillovers: A Reassessment,” *American Economic Review*, 95 (2005), 450–460.

Toivanen, Otto, and Lotta Väänänen, “Education and Invention,” *Review of Economics and Statistics*, 98 (2016), 382–396

Toole, Andrew A., “The Impact of Public Basic Research on Industrial Innovation: Evidence from the Pharmaceutical Industry,” *Research Policy*, 41 (2012), 1–12.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones, “Atypical Combinations and Scientific Impact,” *Science*, 342 (2013), 468–472.

Valero, Anna, and John Van Reenen, “The Economic Impact of Universities: Evidence from across the Globe,” Centre for Economic Performance Discussion Paper, no. 1444, 2016.

Van, der Kloot W., “April 1915: Five Future Nobel Prize–Winners Inaugurate Weapons of Mass Destruction and the Academic–Industrial–Military Complex,” *Notes and Records of the Royal Society*, 58 (2004), 149–160.

Waldinger, F., “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany,” *Review of Economic Studies*, 79 (2012), 838–861.

Waldinger, Fabian, “Quality Matters: the Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany,” *Journal of Political Economy*, 118 (2010), 787–831.

———, “Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge,” *Review of Economics and Statistics*, 98 (2016), 811–831.

Wang, Jian, Reinhilde Veugelers, and Paula Stephan, “Bias against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators,” NBER technical report, 2016.

Weitzman, Martin L., “Recombinant Growth,” *Quarterly Journal of Economics*, 113 (1998), 331–360.

Williams, H. L., “Intellectual Property Rights and Innovation: Evidence from the Human Genome,” *Journal of Political Economy*, 121 (2013), 1–27.

Wuchty, S., B. F. Jones, and B. Uzzi, “The Increasing Dominance of Teams in Production of Knowledge,” *Science*, 316 (2007), 1036–1039.