Dictators, Democracies, and Discoveries: Political Institutions and the Creation of Knowledge*

Sebastian Hager (University of Oxford)
Felix Radde (University of Munich)
Carlo Schwarz (Bocconi University)
Fabian Waldinger (University of Munich)

April 3, 2026

Abstract

We investigate the role of political institutions in shaping global knowledge production. Using a newly assembled global dataset covering universities, scientists, publications, Nobel Prize-winning discoveries, and technological breakthroughs from 1900 to the present, we show that institutional quality is a key determinant of scientific production. Countries with high-quality institutions build larger academic sectors and produce more scientific output, including frontier research and Nobel Prize-winning discoveries. Event-study evidence from sharp institutional improvements and deteriorations confirms that changes in political institutions lead to sustained shifts in scientific production. These effects operate through both a larger research workforce and higher productivity per researcher. Institutions also shape the breadth of inquiry: autocracies concentrate research in a narrower set of fields, achieving excellence in some areas but lacking the broad exploration of ideas that characterizes democracies. Finally, countries with better institutions not only produce more published research but also generate more path-breaking technological discoveries, including those not published in academic journals.

*We are grateful to Davide Cantoni, Edgard Dewitte, Massimo Morelli, Elias Papaioannou, Andrei Shleifer, Adam Szeidl, Guido Tabellini, John van Reenen, Matthias Weigand, and seminar participants at Amsterdam, Bergamo, Bocconi, Cambridge, Lucerne, Munich, MPI Munich, Northwestern, Oxford, Stanford, 3rd Workshop on Funding of Science and Innovation, and the CEPR conference "Public Policies for Innovation" for insightful comments and suggestions. We are grateful to Alessandro Iaria and Lena Greska, who have greatly contributed to the construction of the World of Academia Database, which we use in this paper. We also thank Andrea Amirfeiz, Bianca Batistela, Dario Bonati, Carlotta Colosio, Veljko Kosijer, Saba Kutibashvili, Riccardo Levi, Luiza Marchetti, Ninia Sabadze, Valentina Scaroni, Johannes Scheller, Ascanio Schena, Marie Spörk, Nils Süßenbach, Tomiris Ydyrysbayeva, and the World of Academia RA team for excellent research assistance. E-mail addresses of the authors: Hager: sebastian.hager@economics.ox.ac.uk, Radde: felix.radde@campus.lmu.de, Schwarz: carlo.schwarz@unibocconi.it, Waldinger: fabian.waldinger@econ.lmu.de.

1 Introduction

Economic development ultimately depends on the creation of new knowledge, and countries vary widely in their capacity to produce scientific breakthroughs and frontier innovation (e.g., Romer, 1986, 1990; Mokyr, 1992; Aghion and Howitt, 1992; Jones, 1995). Political institutions shape how states fund, organize, and regulate scientific activity. Because democracies and autocracies may differ systematically in how they perform these functions—ranging from the allocation of research funding to the protection of academic freedom and the governance of research organizations—the effect of political institutions on scientific output is ambiguous.

On the one hand, philosophers and social scientists have argued that democratic institutions are superior at promoting scientific and technological progress (e.g., Popper, 1945; Hayek, 1945; Merton, 1942). When researchers are free to explore new ideas without fear of political or institutional interference, they are more likely to challenge existing paradigms and expand knowledge (Popper, 1945; Polanyi, 1962). On the other hand, successful science often requires coordinated research efforts (e.g., Weinberg, 1961; Price, 1963), and autocracies may be particularly effective at mobilizing resources and directing research toward strategic priorities. This tension between scientific autonomy and state-directed research is central to debates on how political institutions affect scientific progress.

Many historical and contemporary examples highlight this tension. In 1633, the Roman Inquisition placed Galileo Galilei under house arrest for promoting heliocentrism. Three centuries later, in 1933, the Nazi regime expelled roughly 2,000 Jewish and "politically undesirable" academics and burned works by Einstein, Freud, and Marx. More recently, Turkey's Erdoğan administration purged over 6,000 academics and closed several universities. Similarly, since 2025, U.S. science has faced unprecedented pressure under President Trump. While these examples highlight that autocracies suppress scientific activity, some of the most dramatic technological and scientific advances have emerged in authoritarian states. In the 1950s and 1960s, Soviet scientists pioneered space exploration, launching Sputnik, the first satellite, and sending the first human into space. In recent years, Chinese scientists have been advancing the frontier in artificial intelligence, renewable energy, and chemistry. In 2022, China overtook the United States in the Nature Index, a leading measure of high-quality research, and is on course to become the world's largest public funder of science within the next two years.[1]

In this paper, we provide the first comprehensive evidence on the relationship between political institutions and the discovery of ideas at a global scale. To address this question, we assemble a novel dataset covering universities, scientists, publications, Nobel Prize-

[1]For more detail on the Galileo affair, see e.g., Biagioli (1993); on universities in Nazi Germany see e.g., Hartshorne (1937); Grüttner (2024); Waldinger (2010, 2012, 2016); Becker et al. (2024), on universities under Erdoğan, see e.g., Abbott (2016); Kaya (2018); Özatalay (2020), on the impact of the Trump administration on U.S. universities see e.g., Kozlov et al. (2026); Science News Staff (2025b,a); The New York Times (2025). For more detail on the rise of Chinese science, see e.g., Deighton (2026); Ku and Mu (2025); Shepherd and Li (2025); Clarivate Analytics and Chinese Academy of Sciences (2024).

winning discoveries, and technological breakthroughs that spans all countries in the world from 1900 to the present. We document five central findings. First, countries with higher-quality political institutions build larger academic sectors. Second, these countries produce substantially more scientific output, including frontier research and Nobel Prize-winning discoveries. Event-study analyses of major institutional shocks confirm that these relationships are driven by institutional change rather than long-run trends. Third, these effects reflect not only scale but also higher productivity per researcher. Fourth, autocracies channel research into a narrower set of fields than democracies. Fifth, institutional quality increases not only published research but also path-breaking technological discoveries.

For our novel global database, we compile information from several original datasets and established sources. Our first source is the World of Academia Database (Iaria et al., 2022), which provides near-universal coverage of scholars and universities from 1900 to 1969. To provide information on universities up to the present, we complement these data with newly digitized data from the UNESCO World Higher Education Database (IAU, 1985; IAU, 2000; IAU, 2024). To measure scientific output and its impact, we use detailed data on over 200 million publications and their citations from OpenAlex (Priem et al., 2022). Furthermore, we capture breakthrough scientific discoveries by constructing a new dataset of Nobel Prize-winning research. We also obtain information on technological breakthroughs from Asirvatham et al. (2026). We measure political institutions using the V-Dem Liberal Democracy Index (Coppedge et al., 2024), which ranges from 0 (closed autocracy) to 1 (liberal democracy). Finally, we compile historical data on population and GDP per capita from Mitchell (2013) and Gapminder (Gapminder, 2024a,b), as well as conflict data from the Correlates of War Project (Sarkees and Wayman, 2010).

In the first part of the paper, we document a strong positive relationship between institutional quality and the size of a country's academic sector, measured by the number of universities and academics. This relationship is robust to controlling for country and time fixed effects and is not explained by differences in economic development. Our estimates indicate that a 0.1-point increase in the Liberal Democracy Index (LDI), roughly the institutional gap between Italy and Denmark in 2020, is associated with a 24% increase in the number of universities and a 33% increase in the number of academics. These findings highlight the central role of political institutions in shaping the expansion of the academic sector.

In the second part of the paper, we examine whether political institutions influence scientific output. We find that researchers in countries with higher-quality political institutions produce substantially more scientific publications. Two-way fixed effects estimates indicate that a 0.1-point increase in the Liberal Democracy Index is associated with a 13% increase in the number of publications. Given the central importance of high-quality research for scientific and technological progress (e.g., Merton 1957; Iaria et al. 2018), we also analyze the relationship between political institutions and frontier research. We classify a publication as frontier research if its cumulative citation count until

today ranks in the top 10%, top 5%, top 3%, or top 1% of the global citation distribution for publications in the same discipline and year. Our estimates show that a 0.1-point increase in institutional quality is associated with a 11% increase in the number of frontier publications in the top 1%, and similar increases for alternative definitions of frontier research.

While highly-cited publications measure the upper tail of scientific output, they may still not fully capture the most transformative discoveries. To assess whether political institutions also influence truly groundbreaking discoveries, we study effects on Nobel Prize-winning research. Our measure of Nobel Prize-winning research captures the year and country in which future laureates published the discoveries that later earned them the Prize, rather than the year of the award or the laureates' subsequent affiliations. We find that a 0.1-point increase in the Liberal Democracy Index is associated with a 21% increase in the production of Nobel Prize-winning discoveries. The findings on publications, highly-cited publications, and Nobel Prize-winning research suggest a strong link between political institutions and both the quantity and quality of scientific output.

To investigate a potentially causal link between political institutions and scientific production, we study the effect of sudden and major improvements or deteriorations in institutional quality, each defined as a change of at least 0.2 points in the V-Dem Liberal Democracy Index over a five-year window. During the 20^{th} and 21^{st} centuries there were 77 such major improvements, including the democratizations after the fall of the Iron Curtain and Tunisia's transition following the Arab Spring, and 31 major deteriorations, including Franco's fascist dictatorship in Spain, Hitler's totalitarian seizure of power in Germany, Batista's military coup that plunged Cuba into dictatorship and paved the way for decades of communist rule, and Orban's authoritarian turn in Hungary. Using a staggered event-study design, we find that both types of shock have substantial and persistent effects on scientific output: democratic openings spark sustained increases in scientific output while institutional breakdowns trigger sharp and lasting declines in publications and frontier research. The event study estimates also show that countries experiencing sharp institutional changes exhibit no differential pre-trends relative to countries without major changes in institutional quality. The absence of differential pre-trends is consistent with a causal interpretation of the link between political institutions and scientific discoveries.

In the third part of the paper, we distinguish between two margins through which political institutions may shape scientific output: the size of the research sector and productivity per researcher. A central hypothesis is that autocratic regimes may depress scientific production by contracting the scientific workforce; through dismissals, barriers to entry, or by inducing researchers to exit the sector. Alternatively, scientific output in autocracies may decline because researchers who remain in the system become less productive, for example, due to weaker incentives, reduced intellectual autonomy, or adverse selection. To disentangle these channels, we test whether the relationship between political institutions and research output persists once we account for changes in the

number of researchers. We find that, even after controlling for the size of the research workforce, countries with higher-quality political institutions continue to produce more scientific output, although the point estimates are attenuated relative to the baseline specifications. These results imply that political institutions influence scientific output through two key channels: countries with higher-quality institutions not only sustain larger scientific workforces but also exhibit higher productivity per researcher, particularly in producing frontier research. This productivity advantage likely reflects stronger career and publishing incentives, greater intellectual autonomy, and a superior ability to attract and retain highly productive individuals.

In the fourth part of the paper, we study whether political institutions shape not only the quantity and quality of research but also the breadth of scientific inquiry across fields. Philosophers and sociologists of science (e.g., Merton, 1942; Popper, 1945, p. 426; Polanyi, 1962) have argued that authoritarian countries produce narrower research portfolios than democratic countries because democratic regimes allow scientists to freely explore ideas without the risk of political interference. To examine this hypothesis, we quantify the degree of research specialization for each country and year using the distribution of publications across the 256 scientific fields in the OpenAlex data. We measure specialization using concentration ratios, which capture the share of research produced in the k largest fields (e.g., the 25 largest fields) in a country and year. We find that, regardless of the choice of k, frontier research under authoritarian regimes is more narrowly focused. By contrast, countries with liberal democratic institutions produce frontier research across a broader range of fields. The findings suggest that autocracies channel research into a narrower set of fields, achieving excellence in selected areas but lacking the broad exploration of ideas that characterizes democracies.

In the final part of the paper, we examine breakthrough technological discoveries that extend beyond published research, as some scientific findings may not be published in journals or books. In particular, research with military or strategic applications could be classified or concealed for political reasons. Especially, authoritarian regimes may disproportionately divert scientific activity away from publication-focused science toward secretive research. We study whether political institutions affect the creation of path-breaking technological discoveries using data on 20,785 discoveries from Asirvatham et al. (2026). Importantly, the data also include path-breaking discoveries that are not captured in academic publications. These include military innovations, such as the first atomic warhead (USA) and the first ballistic rocket (Nazi Germany); government-led discoveries, such as the Soviet space program; and commercial innovations, such as TikTok (China). The data further cover university-based breakthroughs (e.g., CRISPR-Cas9 gene editing, USA) as well as discoveries by independent inventors, such as the invention of the airplane by the Wright Brothers (USA).

We find that countries with better political institutions produce more technological advances. The estimates imply that a 0.1-point increase in the V-Dem Liberal Democracy

Index is associated with a 3.5% increase in the number of path-breaking technological discoveries. In addition, we show that the overall effect on technological breakthroughs is predominantly driven by discoveries in universities and industry, with somewhat smaller effects on government-led discoveries. In contrast, autocratic regimes disproportionately promote discoveries in military technologies, even though the difference to democracies is not significantly different from 0. Finally, political institutions have little discernible impact on independent inventors, whose serendipitous discoveries may rely less on formal institutional quality.

Prior work in economics has largely focused on specific episodes of autocratic interference in science and innovation. Some studies highlight potential positive effects of autocracies, such as the role of government demand for surveillance technology in stimulating AI research in China (Beraja et al., 2023; Beraja, Yang, and Yuchtman, 2023), or the decline of Russian science after the collapse of the Soviet regime (Ganguli, 2014; Borjas and Doran, 2015). Other contributions emphasize the harmful effects of autocracies on science and universities, studying episodes such as the Spanish and Roman Inquisition (Cox and Figueroa, 2025; Dewitte et al., 2024), book censorship under China's Qing dynasty (Bai et al., 2024), and the dismissal of scientists during the Nazi period (Waldinger, 2010, 2012, 2016; Moser et al., 2014; Becker et al., 2024). While not studying the effect of political institutions on science and innovation, two recent papers have correlated measures of academic freedom (rather than political institutions) with patenting and publishing at the country level (Audretsch et al., 2024; Fernandez et al., 2024).[2] In contrast to these earlier studies, our paper provides the first global and long-run analysis of how political institutions shape scientific and technological progress, assembling new data to cover nearly all countries worldwide throughout the 20$^{\text{th}}$ century and beyond.

Our paper also contributes to the extensive literature examining the relationship between political institutions and economic growth. A large body of theoretical and empirical literature indicates that inclusive political institutions lead to sustained economic prosperity (North, 1990; Rodrik et al., 2004; Acemoglu et al., 2005; Acemoglu and Robinson, 2013; Acemoglu et al., 2019; Funke et al., 2023). Another literature has highlighted that general education is vital for economic growth (e.g., Glaeser et al., 2004; Gennaioli et al., 2013), that countries with a higher proportion of engineering majors grow faster (e.g., Murphy et al., 1991; Maloney and Valencia Caicedo, 2022), that regions with more and better funded universities grow faster (e.g., Valero and Van Reenen, 2019; Kantor and Whalley, 2014), and that upper-tail human capital was an important driver of the industrial revolution (e.g., Mokyr, 2002, 2005; Squicciarini and Voigtländer, 2015). Our results highlight a complementary channel through which political institutions foster economic development: they expand the stock of human capital at the very top of the skills distribution and enhance the production of ideas and path-breaking discoveries.

[2]Relatedly, Gao et al. (2017); Ahmed et al. (2024) report no significant effect of democratization on U.S. patents and the OECD Global Innovation Index.

2 Data

Our analysis is based on a unique combination of several large databases of the academic sector covering all countries of the world from 1900 until today. This allows us to reconstruct information on the number of universities, the number of academics, detailed academic output measures, and technological breakthroughs on a worldwide scale, covering the entire 20$^{\text{th}}$ century and the first two decades of the 21$^{\text{st}}$ century.

2.1 Number of Universities

To reconstruct the number of universities in each country over the entire 20$^{\text{th}}$ century and beyond, we rely on two data sources. The first is the *World of Academia Database* (Iaria et al., 2022), a 12-year data collection effort, which captures the near universe of worldwide universities for the years 1900, 1914, 1925, 1938, 1956, and 1969.[3] It is the most comprehensive listing of worldwide universities, and of the academics working in these institutions for the first seven decades of the twentieth century.[4]

As the publication of the data underlying the *World of Academia Database* ceased after 1969, we obtain information on the number of universities by country from three publications by the International Association of Universities and UNESCO: the *World List of Universities* (IAU, 1985) for 1985, the *World List of Universities* (IAU, 2000) for 2000, and the *UNESCO World Higher Education Database* (IAU, 2024) for 2020.

By integrating these two sources, we reconstruct the most comprehensive census of worldwide universities to date. The data span the period from 1900 to 2020 and report the number of universities in each country for the following points in time: 1900, 1914, 1925, 1938, 1956, 1969, 1985, 2000, and 2020. A key advantage of the new dataset is that both sources provide contemporaneous snapshots of the university sector; hence, we avoid conditioning the analysis on the selected sample of universities surviving until today.

Because many universities initially operate as small or specialized institutions, they may not appear in contemporaneous listings if founded shortly before the publication of a historical source. To construct accurate university counts by country and year, we exploit information on founding dates across multiple historical snapshots of the university landscape. Specifically, we use information on universities' founding years from later cross-sections to infer how many institutions must have existed in earlier periods, thereby

[3]Because the number of universities expanded rapidly over time, the data for the 1950s and 1960s were released in multiple installments. We assign each decade's observation to the publication year of its final installment—1956 for the 1950s and 1969 for the 1960s.

[4]As documented in Iaria et al. (2022), the World of Academia database is constructed from the historical editions of the Minerva Jahrbuch der Gelehrten Welt. Minerva's editors sought near-universal coverage by soliciting information from education ministries, university administrations, and individual academics. Contemporary observers also commented on Minerva's breadth: for example, a Nature article contrasted it with the French Annuaire Général des Universités, noting that while the two served similar purposes, the French source was "very much less exhaustive" (Nature, 1930). Independent benchmarking further shows that Minerva captured a large share of the global university landscape (Iaria et al., 2022).

recovering universities that are absent from contemporaneous records (see Appendix A.1 for details).[5]

For the main results, we focus on "main" universities, i.e., traditional universities (e.g., Harvard, Oxford, or the Sorbonne) and technical universities (e.g., MIT, École Polytechnique, or the Technical University of Munich) but exclude more specialized institutions such as nursing schools, art institutions, and conservatories.[6] We illustrate the geographic distribution of universities as of 1900, 1956, and 2020 in Figure 1.

2.2 Number of Academics

We reconstruct the number of academics in "main" universities in each country from the microdata in the *World of Academia Database* (Iaria et al., 2022). The database contains faculty rosters for the near-universe of universities in all countries of the world for six cross-sections covering academics in 1900, 1914, 1925, 1938, 1956, and 1969. We consider all academics in a university, i.e., full professors, associate professors, assistant professors, and honorary professors.

As outlined above, for the years until 1969, the database provides the most comprehensive data on worldwide academics. As neither the *World List of Universities* nor the *World Higher Education Database* contain microdata on academics, we cannot reconstruct the number of academics for the post-1969 period.[7]

2.3 Publications

We construct data on the number of publications for each country of the world in five-year intervals from 1900 until 2020 (i.e., 1900, 1905, 1910, ..., 2015, 2020) from *OpenAlex* (Priem et al., 2022). *OpenAlex* is an open-source index of scholarly publications that aggregates metadata on over 200 million academic publications, including information on authorship, institutional affiliation, publication year, citation counts, as well as discipline and field classifications.

We assign publications to countries based on authors' institutional affiliations.[8] Because *OpenAlex* maps affiliations to contemporary country borders, output from countries that experienced territorial changes would be systematically misclassified. For instance, a paper published in 1926 in Breslau (now Wrocław) would be attributed to Poland rather than to Germany, its historically correct country at the time of publication. To correct such misclassification, we use yearly shapefiles of historical borders (Schvitz et al., 2022) and

[5]Our results are unchanged when we do not apply this adjustment.

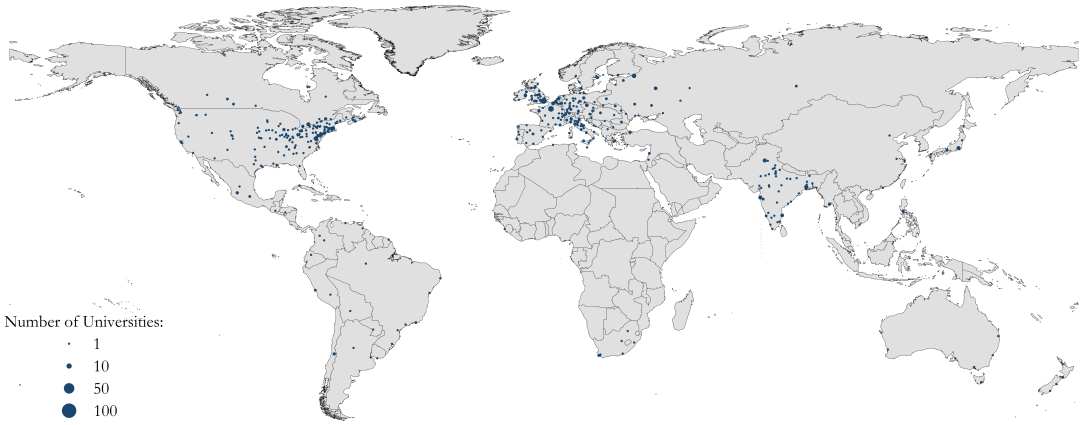
[6]We identify regular universities and technical universities with a GPT-based algorithm (see Appendix A.1 for details). The results are robust to using alternative definitions of universities.

[7]We are not aware of comprehensive microdata on all academics in the world in all universities and disciplines for the post-1969 period.

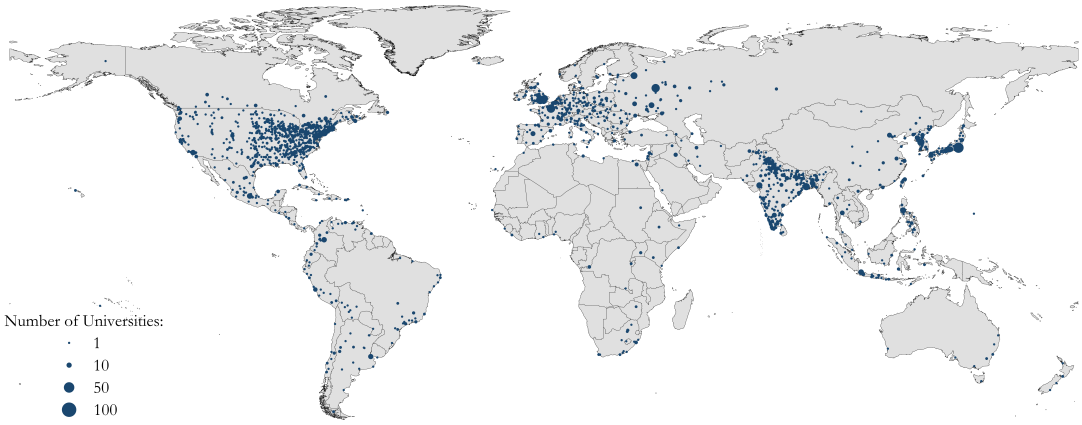
[8]Note, some publications in the OpenAlex data do not report an affiliation and can therefore not be assigned to a country. We focus our analysis on publications that can be assigned to a country.

Figure 1: Universities by City Over Time

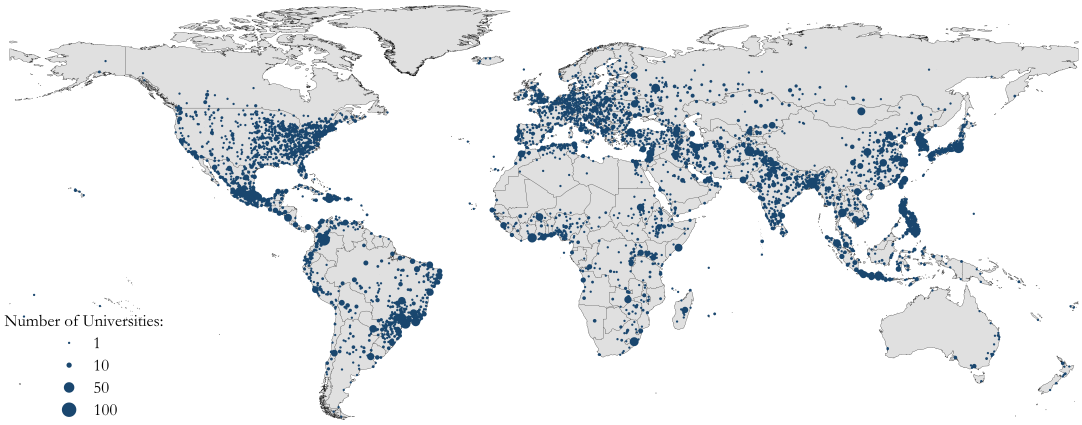
(a) 1900



(b) 1956



(c) 2020



Notes: The maps show the total number of universities (regular universities and technical universities) by city for the years 1900, 1956, and 2020. Country boundaries correspond to historical borders in each period. Data have been collected by the authors from the World of Academia Database (Iaria et al., 2022), the *World List of Universities* (IAU, 1985; IAU, 2000), and the *UNESCO World Higher Education Database* (IAU, 2024) (see Section 2.1 for details).

affiliation coordinates reported in OpenAlex to assign historically accurate countries to each publication.[9]

We measure original research output using the number of journal articles, book chapters, and books recorded in OpenAlex for each country and time period. To capture the entire scientific production in a country, we aggregate publications within a ±2-year window around each observation, i.e., we count publications from 2018 to 2022 for the 2020 observation. This smoothing procedure mitigates measurement error, particularly for small countries and earlier periods with limited publication volumes. The results are robust to alternative definitions, including a ±1-year window or counting publications in year t only.

It is important to note that *OpenAlex* records all scientific output produced within a country, regardless of the institutional affiliation of the authors. Our measure, therefore, encompasses publications not only by university academics but also by researchers in independent research institutes (e.g., the Academy of Sciences of the Soviet Union, CNRS, or the Max Planck Institutes), or by scientists working in the private sector.

We also use citation data from *OpenAlex* to identify highly-cited publications, specifically those in the top 10%, top 5%, top 3%, or top 1% of the discipline and year-level citation distribution.[10] Specifically, a publication is classified as highly-cited if its cumulative citation count until today falls within the corresponding percentile of the citation distribution of all papers published in the same discipline and within the same ±2-year window. For example, we count the number of publications which were produced by mathematicians in country c in a ±2-year window around year t whose citation count until today falls in the top 1 percent of the worldwide citation distribution of mathematics publications published in that time window.

2.4 Nobel Prize Winning Discoveries

We assemble data on Nobel Prize-winning research, identifying the year and country in which future laureates published the work that ultimately earned them the prize. For prizes awarded between 1900 and 2016, we draw on data from Li et al. (2019), which reports Nobel-winning publications including OpenAlex identifiers. We link these records to OpenAlex to recover the historically correct country of publication and systematically hand-check the linked publications to ensure a correct match.[11] For prizes awarded between

[9] In a small number of cases, the historical shapefiles of Schvitz et al. (2022) are imprecise along coastlines, so that some affiliation coordinates fall just offshore. We manually inspect these observations and assign them to the correct country. In addition, for the period 1949-1989 we manually assign publications from Berlin to either East or West Berlin (the GDR or the Federal Republic of Germany), as appropriate.

[10] We use the "field" variable in OpenAlex to classify publications into 26 disciplines (e.g., mathematics, physics, or humanities).

[11] Information on a few pre-2016 laureates is missing in the Li et al. (2019) data. For these cases, we use data from Jones and Weinberg (2011) to add information on the year and country of producing their Nobel Prize-winning research.

2017 and 2024, we extend the dataset by manually collecting the Nobel Foundation's official scientific background documents. From these documents, we extract all references to publications authored by the laureates. We define Nobel Prize-winning publications as the complete set of these references.

We assign a total weight of one to each Nobel Prize (e.g., the 1921 Nobel Prize in Physics) and distribute this weight across Nobel-winning contributions according to the number of laureates, their prize-relevant publications, and the institutional affiliations at the time those publications were produced. For example, Albert Einstein received the 1921 Nobel Prize when he was working in Germany. The award recognized "his services to Theoretical Physics" (Nobel Foundation, 2025c), and especially his 1905 *Annalen der Physik* article, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt," written while he was employed at the Swiss Patent Office. Because Einstein was the sole recipient of the 1921 prize, we allocate the full weight of 1 to Switzerland in 1905 for the 1921 Nobel Prize in physics (Table 1, column 9).

As another illustration, consider the 1958 Physics Nobel Prize, which was awarded to Pavel Cherenkov, Ilya Frank, and Igor Tamm "for the discovery and interpretation of the Cherenkov effect" (Nobel Foundation, 2025d). We assign each laureate one-third of the prize (Table 1, column 2). Cherenkov's Nobel-winning research was published in 1934, while Frank and Tamm's research appeared in 1937; accordingly, we allocate 0.33 units of the prize to the Soviet Union in 1934 and 0.66 units to 1937 (column 9).

Our final example is the 2020 Nobel Prize in Chemistry, awarded to Emmanuelle Charpentier and Jennifer Doudna for the "development of a method for genome editing" using CRISPR-Cas9 (Nobel Foundation, 2025b). We assign each laureate a prize weight of 0.5 (column 2). Charpentier's contribution rests on two key papers (within-laureate paper weights of 0.5 each; column 6), while Doudna's rests on one paper (within-laureate weight of 1; column 6). Country weights follow the affiliations listed in the publications: Charpentier's first paper lists Austria and Sweden (0.5 each; column 8), her second lists Sweden only (weight 1, column 8), and Doudna lists only U.S. affiliations (weight 1; column 8). Multiplying these components yields the following weights: 0.125 each for Austria and Sweden in 2011, 0.25 for Sweden in 2012, and 0.5 for the United States in 2012 (column 9).

We apply this procedure to all Nobel Prizes. This ensures that our measure reflects the timing and geography of the underlying Nobel-winning research, rather than the year of the award or the laureates' subsequent affiliations. After constructing the final country-level weights for each Nobel Prize, we aggregate them to the country-year level and sum over a ± 2-year window around year t. Our new measure of Nobel Prize-winning discoveries traces the shifting geography of Nobel-winning research. In the early 1900s, France was the global leader, closely followed by Germany. By the mid-1920s, Germany had overtaken France, with the United States in second place. After 1950, the center of gravity in science shifted across the Atlantic: the United States became the dominant producer of Nobel-winning research and has maintained that position ever since.

Table 1: Nobel Prize-Winning Research Weighting Examples

(1) Laureate	(2) Laureate Weight	(3) Publication	(4) Journal	(5) Publication year	(6) Within Laureate paper weight	(7) Country at time of publication	(8) Within paper-Laureate country weight	(9) Weighted Nobel winning research: (2) × (6) × (8)
1921 Nobel Prize in Physics								
Einstein	1	Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt	Annalen der Physik	1905	1	Switzerland	1	1
1958 Nobel Prize in Physics								
Cherenkov	0.33	Visible emission of pure liquids under the action of gamma radiation	Comptes Rendus (Dokl.) Acad.Sci. USSR	1934	1	Soviet Union	1	0.33
Frank	0.33	Coherent visible radiation from fast electrons passing through matter	Comptes Rendus (Dokl.) Acad.Sci. USSR	1937	1	Soviet Union	1	0.33
Tamm	0.33	Coherent visible radiation from fast electrons passing through matter	Comptes Rendus (Dokl.) Acad.Sci. USSR	1937	1	Soviet Union	1	0.33
2020 Nobel Prize in Chemistry								
Charpentier	0.5	CRISPR RNA maturation by trans-encoded small RNA and host factor Rnase III	Nature	2011	0.5	Austria	0.5	0.125
						Sweden	0.5	0.125
		A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity	Science	2012	0.5	Sweden	1	0.25
Doudna	0.5	A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity	Science	2012	1	USA	1	0.5

Notes: The table illustrates how we assign Nobel Prize-winning discoveries to countries and years (see Section 2.4 for details).

2.5 Path-Breaking Technological Discoveries

We complement publication-based measures of scientific output by analyzing a broad set of path-breaking technological discoveries, including those that did not appear in academic journals or books (e.g., classified military research). To measure such discoveries, we use data from Asirvatham et al. (2026). These data are based on the English-language Wikipedia corpus, which they parse through a GPT-based extraction pipeline. The procedure filters out historically significant technologies, removes conceptual duplicates, and then enriches each discovery with structured attributes.[12] For each discovery, the procedure extracts the year, the country, the institution, and the institutional type (university, government, military, corporate, or independent inventor).

We use the microdata to construct country-year measures of technological breakthroughs for the period 1900–2020. For each country and period, we count the number of inventions originating in that country and aggregate them using the same ±2-year window around year t as for our other outcome measures.

2.6 Political Institutions

We measure political institutions with the Varieties of Democracy (V-Dem) Liberal Democracy Index (LDI) (Coppedge et al., 2024). The LDI is a widely used index that measures whether a country can be considered a liberal democracy. The LDI captures two main dimensions: first, whether a country allows for free, fair, and reoccurring elections ("electoral democracy") and, second, whether a country's institutions place "constraints on the executive by the legislature and the judiciary, and the rule of law ensuring respect for civil liberties" (Nord et al., 2025). The index is based on assessments from over 4,200 country experts, predominantly academics, whose responses are aggregated using a Bayesian item-response theory model to ensure cross-national and intertemporal comparability (Coppedge et al., 2019). The index ranges from 0 to 1, where 1 is a full liberal democracy. We visualize the V-Dem Liberal Democracy index for all countries of the world for the years 1900, 1960, and 2000 in Figure 2.[13]

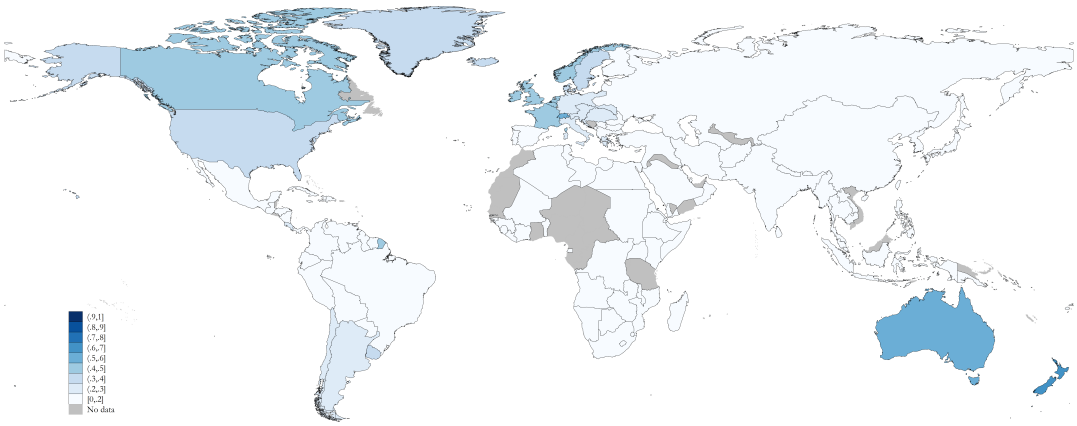
A key advantage of V-Dem is that it reports institutional quality both for sovereign states and also for semi-sovereign territories and colonies (e.g., India prior to independence in 1947). This broader coverage enables us to study the long-run effect of institutions on science, including in territories that became independent only later in the twentieth century. Compared with alternatives such as Polity (Marshall and Gurr, 2020), which restricts coverage to sovereign states, V-Dem provides substantially broader geographic coverage across the 20$^{\text{th}}$ century. For instance, for years before 1950, Polity includes only about half as many countries or territories as V-Dem (Appendix Figure A.2). For the

[12]The data contain additional duplicates that we filter out using a further de-duplication step (see Appendix A.2).

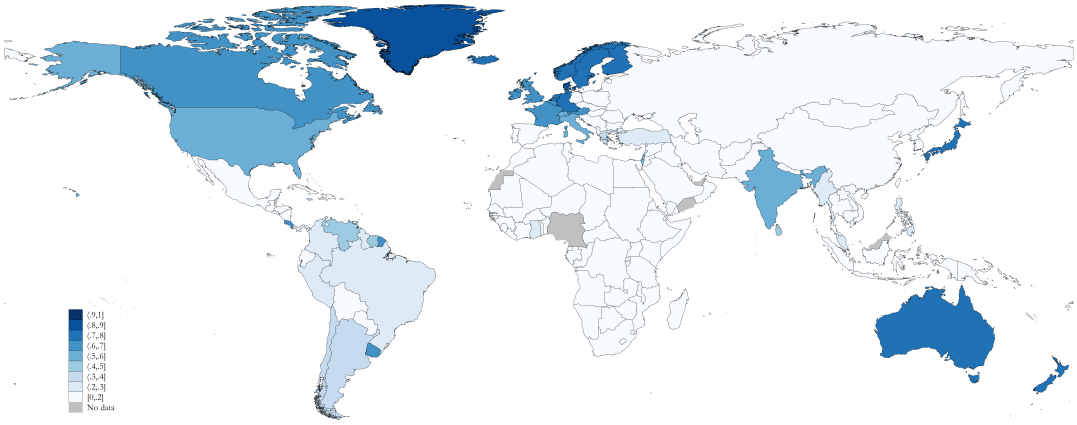
[13]Note that there were only 114 countries and semi-independent territories in 1900. By 2020, the number of countries had increased to 195.

Figure 2: Political Institutions Over Time

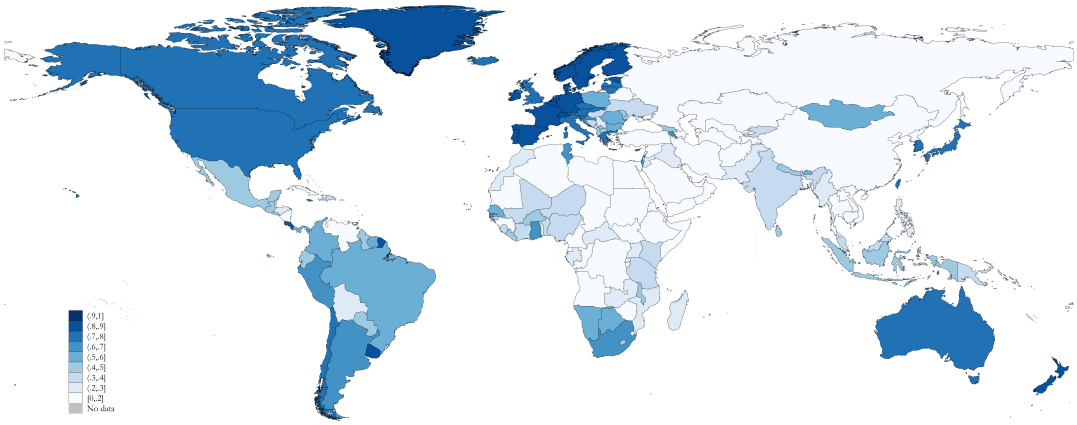
(a) 1900



(b) 1960



(c) 2020



Notes: The maps show the V-Dem Liberal Democracy Index for the years 1900, 1960, and 2020. Country boundaries correspond to historical borders in each period. Territories shaded in grey indicate that no value of the Liberal Democracy Index is available from V-Dem for the respective year.

countries covered in both sources, the correlation between the two measures of institutional quality is 0.84 (Spearman rank correlation = 0.85).[14]

2.7 Other Country-Level Data

We also collect data on population and GDP per capita to control for underlying country characteristics that may influence scientific activity. Country-level population figures by year come from International Historical Statistics (IHS) (Mitchell, 2013). Unlike alternative population figures, such as Maddison, the IHS data compile data from historical censuses. Most importantly for our long-run analysis, IHS reports populations according to historically accurate country borders, not modern borders.[15] For most countries, International Historical Statistics provides data for census years. We linearly interpolate between these years to construct annual estimates. To capture border changes that coincide with large changes in population, we do not extrapolate across years when a country experienced a border change. Because the IHS data were discontinued after 2013 and do not report pre-2013 data for a small number of countries, we supplement the population data with population estimates from Gapminder (2024b).

We also collect annual country-level data on GDP per capita from Gapminder (2024a), a harmonized dataset that integrates economic indicators from multiple sources, including the World Bank, the Maddison Project, the Penn World Table, and Gapminder's historical GDP reconstructions (primarily compiled by Lindgren, 2011).

2.8 Final Dataset

For the analysis, we merge all datasets described above into a country-level panel spanning 1900 to 2020, the most comprehensive global record of scientific activity and institutions assembled to date. We aggregate all variables into five-year intervals, covering nearly the entire universe of sovereign states, semi-sovereign territories, and colonies over the past 120 years.[16] The data contain harmonized measures of universities, academics, publications, highly-cited research, Nobel Prize-winning discoveries, technological breakthroughs, political institutions, and economic fundamentals, all mapped to historically accurate borders. Summary statistics are reported in Table 2.

[14] A number of recent economics papers have used V-Dem scores to measure institutional quality (e.g., Funke et al., 2023; Acemoglu et al., 2024).

[15] For example, Maddison reports Austria's population in 1910 as 6,571,595, whereas the true population of the Austrian part of the Habsburg Empire was 28,572,000.

[16] The only exceptions are the data on the number of academics, which rely on World of Academia snapshots for 1900, 1914, 1925, 1938, 1956, and 1969, and the data on the number of universities, which use the same World of Academia years supplemented with information for 1985, 2000, and 2020 from the World List of Universities and the World Higher Education Database.

Table 2: Summary Statistics at the Country-Time Level

Variable	Mean	Std. Dev.	Obs.
Political Institutions			
Liberal Democracy Index	0.25	0.25	3,774
Size of the Academic Sector			
Universities	28.97	116.69	1,343
Academics at Universities	670.98	4,432.34	774
Scientific Discoveries			
Publications	20,074.35	137,034.24	3,774
Highly-Cited Publications (Top 10%)	1,918.35	16,245.97	3,774
Highly-Cited Publications (Top 5%)	957.08	8,596.30	3,774
Highly-Cited Publications (Top 3%)	574.42	5,370.64	3,774
Highly-Cited Publications (Top 1%)	191.83	1,933.52	3,774
Weighted Nobel Prize-Winning Discoveries	0.08	0.66	3,774
Technological Breakthroughs			
Technological Discoveries	5.50	48.41	3,774
Technological Discoveries by Universities	1.28	11.95	3,774
Technological Discoveries by Government	0.72	4.82	3,774
Technological Discoveries by Military	0.24	2.63	3,774
Technological Discoveries by Industry	2.88	28.50	3,774
Technological Discoveries by Independent Inventors	0.26	2.07	3,774

Notes: The table reports summary statistics at the country-year level. Data on most variables are available for all countries of the world in 5-yearly intervals for the period 1900–2020. Data on the number of universities are available for the following years: 1900, 1914, 1925, 1938, 1956, 1969, 1985, 2000, and 2020. Data on the number of academics are available for the following years: 1900, 1914, 1925, 1938, 1956, and 1969.

3 Political Institutions and the Size of the Academic Sector

In the first part of our analysis, we examine how political institutions relate to the size of a country's academic sector. We focus on two key outcomes: the number of universities and the size of the academic workforce employed in universities.

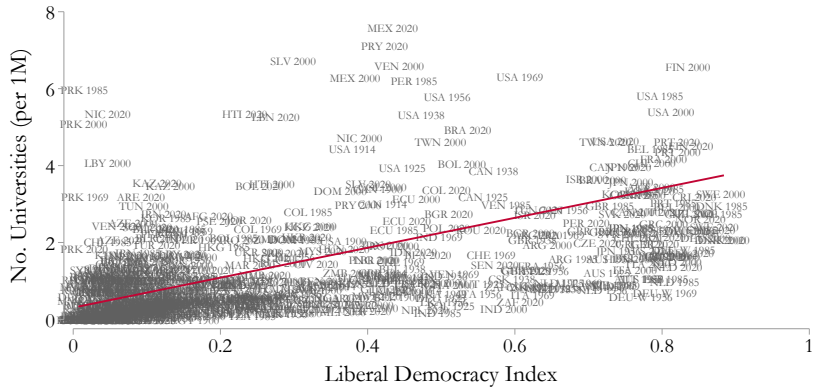
3.1 Number of Universities (1900–2020)

We first use our newly compiled dataset to examine the relationship between political institutions and the number of universities. The dataset reports the number of universities for all countries in 1900, 1914, 1925, 1938, 1956, 1969, 1985, 2000, and 2020. We first plot the number of universities in each country and year (e.g., the United States in 2020 or France in 1938) against its V-Dem Liberal Democracy Index (Figure 3a).

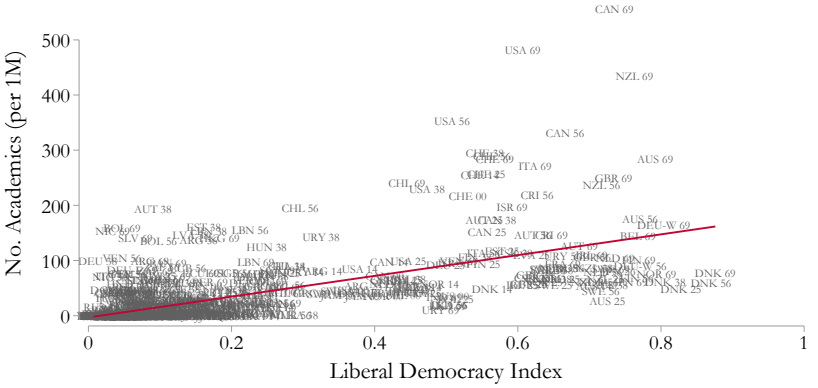
The scatter plot shows a strong positive association between institutional quality and the number of universities. The fitted regression line indicates that a 0.1 point increase in the V-Dem Liberal Democracy Index (roughly equivalent to the institutional quality gap between Italy and Denmark in 2020) is associated with an increase of 0.39 in the number

Figure 3: Political Institutions and the Size of the Academic Sector

(a) Institutions and the Number of Universities



(b) Institutions and the Number of Academics



Notes: Panel (a) plots the number of universities per one million inhabitants (vertical axis) against the Liberal Democracy Index (horizontal axis) for country-year observations from 1900 to 2020. Panel (b) plots the number of academics per one million inhabitants (vertical axis) against the Liberal Democracy Index (horizontal axis) for country-year observations from 1900 to 1969.

of universities per 1 million inhabitants. This corresponds to a 32 percent increase relative to the mean number of universities across all countries and years in the sample.

While the scatter plot suggests a strong positive relationship between political institutions and the number of universities, the relationship may be spurious because of unobserved factors. For example, there may be fixed country characteristics that are correlated with both institutional quality and the number of universities. Similarly, common time shocks and secular trends that shape university expansion may co-move with over-time institutional change. To address these concerns, we exploit within-country and over-time variation by estimating the following two-way fixed effects regression:

$$\text{Universities}_{ct} = \phi \cdot \text{Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \tag{1}$$

This specification exploits within-country variation over time, controlling for unobserved time-invariant country characteristics with country fixed effects and for common global shocks with year fixed effects. The dependent variable, Universities_{ct}, is the number of

universities in country c and time t. The main explanatory variable, Institutional Quality$_{ct}$, is the V-Dem Liberal Democracy Index, which ranges between 0 and 1. X_{ct} is a vector of control variables that includes population and GDP per capita. The controls account for cross-country differences in economic development that may influence the size of the academic sector. We lag the measure of political institutions and the control variables by one year to allow for delays between changes in political institutions and other economic factors and subsequent university openings or closures. We cluster standard errors at the country level to allow for arbitrary correlation of regression residuals within countries over time.

The estimates indicate that institutional quality is positively associated with the number of universities per capita (Table 3, Panel A, column 1). Adding year fixed effects absorbs the secular expansion of university systems over the 20^{th} century, which may be correlated with general improvements in institutional quality, but only has a modest impact on the estimates (column 2). When we additionally control for country fixed effects, so identification comes from within-country changes in institutional quality over time, the estimate declines somewhat but remains statistically significant at the 1% level (column 3). The country fixed effects absorb time-invariant country characteristics that simultaneously influence institutional quality and the university sector, such as colonial legacies (e.g., Acemoglu et al., 2001) or historical religious traditions (e.g., Cantoni and Yuchtman, 2014). In the last column, we additionally control for GDP per capita to address the concern that richer countries may fund more universities and, at the same time, have higher-quality institutions.[17] The estimated coefficient of 69.3 implies that an increase in the Liberal Democracy Index of 0.1 (roughly equivalent to the institutional quality gap between Italy and Denmark in 2020) is associated with a 24 percent increase in the number of universities.

The results are robust to alternative transformations of the dependent variable. Using the logarithm of the number of universities, which drops country-year observations with zero universities, yields similar estimates, as does applying the inverse hyperbolic sine transformation (Appendix Table B.1). The findings are also robust to counting all higher-education institutions rather than only regular and technical universities (Appendix Table B.2). Finally, results are similar when we do not backfill university counts (see Section 2.1) for institutions founded, but not yet recorded in contemporaneous listings, and when we vary the length of the backfilling window (Appendix Table B.3).

3.2 Academic Workforce (1900–1969)

Next, we study the relationship between political institutions and the size of the academic workforce. The number of universities and the size of the academic workforce do not necessarily move in tandem. Authoritarian regimes, in particular, may seek to centralize

[17] Note, however, that GDP per capita may be an endogenous control variable in this context.

Table 3: Political Institutions and the Size of the Academic Sector

	(1)	(2)	(3)	(4)
Panel A: Dep. Var.: Universities (1900–2020)				
Institutional Quality	94.93**	81.90*	74.87***	69.29***
	(38.85)	(43.99)	(28.42)	(25.75)
Observations	1,343	1,343	1,343	1,343
R^2	0.255	0.259	0.692	0.693
Dependent Variable Mean	28.97	28.97	28.97	28.97
Panel B: Dep. Var.: Academics at Universities (1900–1969)				
Institutional Quality	4432.2**	4209.1**	2521.3***	2163.7**
	(1984.2)	(1877.8)	(892.1)	(918.0)
Observations	774	774	774	774
R^2	0.146	0.151	0.665	0.671
Dependent Variable Mean	670.98	670.98	670.98	670.98
Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: Panel A reports estimates of Equation (1) with the number of universities in country c and year t as the dependent variable. Panel B reports estimates of Equation (2) with the number of academics as the dependent variable. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for population. In column (2), we additionally include time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

control by maintaining fewer but larger universities that are easier to monitor and regulate. For example, between 1949 and 1961, China reorganized its higher education system by drastically reducing the number of institutions, and eliminating private and missionary universities to reinforce state authority over the sector (Hsu, 1964).

We measure the academic workforce by the number of university academics in a country and year, drawing on data from the *World of Academia Database*. The database provides faculty rosters for universities worldwide for the years 1900, 1914, 1925, 1938, 1956, and 1969. As a first step, we illustrate the relationship between political institutions and the size of the academic workforce in a scatter plot (Figure 3b). The figure reveals a strong positive association between institutional quality and the number of academics, mirroring the results for the number of universities.

We next examine whether the positive association between institutional quality and the number of academics also holds when exploiting within-country variation over time, by estimating the following two-way fixed effects regression:

$$\text{Academics}_{ct} = \phi \cdot \text{Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \tag{2}$$

where the dependent variable, Academics_{ct}, is the number of academics in country c and

year t. As above, Institutional Quality$_{ct}$ is the V-Dem Liberal Democracy Index, which ranges between 0 and 1. We also add the same control variables as in Equation (1).

The results confirm that institutional quality is positively associated with the number of academics (Table 3, Panel B). This relationship holds when we control for time fixed effects, country fixed effects, and GDP per capita (column 4). The estimated coefficient implies that a 0.1 increase in the Liberal Democracy Index is associated with a 33 percent increase in the number of academics. These findings indicate that countries with better institutions not only establish more universities but also increase the academic workforce.

As for the results on universities, the findings are robust to alternative transformations of the dependent variable. Using the logarithm of the number of academics, which drops country-year observations with zero academics, yields similar estimates, as does applying the inverse hyperbolic sine transformation (Appendix Table B.1).

4 Political Institutions and Scientific Output

In the main part of our analysis, we study how political institutions influence a country's capacity to generate scientific output and discoveries. As described above, our data capture scientific output published in academic journals and books not only by university researchers, but also by scientists working in research institutes and in the private sector.

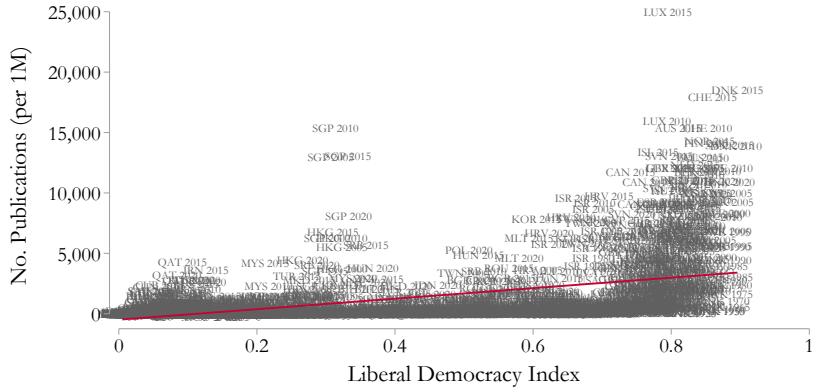
As a first step, we illustrate the relationship between political institutions and scientific output using a scatter plot (Figure 4). Each observation corresponds to a country c in period t. Because publication counts are small for some countries, especially in the early twentieth century, we aggregate output into 5-year intervals. Accordingly, output in period t includes publications in a $\pm$2-year window around year t. For comparability across countries and over time, we express scientific output per capita by normalizing publication counts by population.

We document a strong positive association between institutional quality and the number of scientific publications (Figure 4, panel a). The fitted regression line indicates that a 0.1 increase in the V-Dem Liberal Democracy index is associated with an increase of 435 publications per one million inhabitants. This corresponds to a 62% increase relative to the mean number of publications across all countries and time periods.

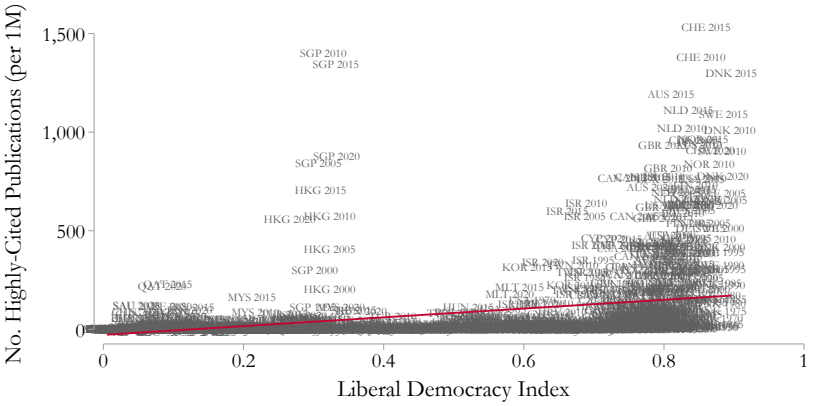
Next, we investigate the relationship between institutional quality and the number of highly-cited publications (Figure 4, panel b). We classify a publication as highly cited if its citation count until today falls within the top five percent of the worldwide citation distribution among publications in the same discipline and time period. The regression line indicates that a 0.1 increase in the V-Dem Liberal Democracy Index is associated with an increase of 22 highly-cited publications per one million inhabitants, i.e., an increase of 68%.

Figure 4: Political Institutions and Publications

(a) Publications



(b) Highly-Cited Publications (Top 5%)



Notes: Panel (a) plots the number of publications per one million inhabitants (vertical axis) against the Liberal Democracy Index (horizontal axis) for country c and time t. Publications are measured in a $\pm$2-year window around time t, which ranges from 1900 to 2020 in 5-yearly intervals. Panel (b) plots the number of highly-cited publications per one million inhabitants (vertical axis) against the Liberal Democracy Index (horizontal axis). Highly-cited publications (i.e., publications in the top 5% in the discipline and year-level citation distribution) are measured in a $\pm$2-year window around time t.

4.1 Institutions and Scientific Output: Regressions Controlling for Country and Time Confounders

While the scatter plots suggest a strong relationship between political institutions and scientific output, they may reflect a spurious correlation because they do not account for unobserved factors that are correlated with institutional quality and publication activity. For example, countries with higher-quality institutions may exhibit characteristics that increase their research capacity, such as a historical scientific tradition or enduring cultural attitudes toward scholarship. Moreover, over time changes in institutional quality may be correlated with general increases in the number of publications. To address these concerns, we estimate two-way fixed effects regressions that include country and time fixed effects as well as time-varying control variables. This specification absorbs time-invariant cross-country differences and common global trends, and relates scientific output

to within-country changes in institutional quality over time.

Number of Publications

We study the relationship between political institutions and the number of publications by estimating the following regression:

$$\text{asinh(Publications)}_{ct} = \phi \cdot \text{ Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \quad (3)$$

where the dependent variable, $\text{asinh(Publications)}_{ct}$, is the inverse hyperbolic sine of the number of publications in country c and time t. As above, we measure publications in a $\pm$2-year window around year t. The cross-country publication distribution has a long right tail. For example, OpenAlex reports more than three million publications by U.S. scientists in the $\pm$2-year window around 2015. We therefore use the inverse hyperbolic sine transformation of the number of publications. Unlike the natural logarithm, the inverse hyperbolic sine transformation remains defined at zero. This is important as in a few countries in the early part of the twentieth century (e.g, Trinidad and Tobago in 1905) researchers did not publish scientific papers covered by the OpenAlex database.[18] $\text{Institutional Quality}_{ct}$ is the V-Dem Liberal Democracy Index. Most importantly, we include country and time fixed effects in the regression. We also add control variables (X_{ct}), including the inverse hyperbolic sine of the population and GDP per capita. We lag the Liberal Democracy Index and the control variables by one year to account for delays between changes in political institutions and economic factors and subsequent scientific production. We cluster standard errors at the country level to allow for arbitrary correlation of regression residuals within countries over time.

We find that countries with higher-quality political institutions produce more scientific output (Table 4, column 1). The estimates shrink slightly if we include year fixed effects, which account for the overall expansion of scientific publishing over the last 120 years (column 2). We additionally include country fixed effects (column 3) and a control for a country's GDP per capita (column 4). The coefficients remain significant at the 1%-level in all specifications. The estimate of 1.217 (column 4) implies that a 0.1 increase in the V-Dem Liberal Democracy Index, roughly the institutional gap between Italy and Denmark in 2020, is associated with a 13% increase in the number of scientific publications.

Highly-Cited Publications

As high-quality research is of particular importance for scientific and technological progress, we also investigate whether political institutions affect frontier research. We classify publications as frontier research if their citation count until today falls within the top 10%, top 5%, top 3%, or top 1% of the worldwide citation distribution among publications in

[18] The results are robust to using the log transformation and thus dropping country-year observations with 0 publications (see Appendix Table C.1).

Table 4: Political Institutions and Publications

	Dep. Var.: Asinh Publications (1900–2020)			
	(1)	(2)	(3)	(4)
Institutional Quality	7.762***	5.961***	1.317***	1.217***
	(0.279)	(0.331)	(0.312)	(0.305)
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes
Observations	3,774	3,774	3,774	3,774
R^2	0.726	0.823	0.946	0.947
Dependent Variable Mean	5.30	5.30	5.30	5.30

Notes: The table reports estimates of Equation (3), where the dependent variable is the inverse hyperbolic sine of the number of publications in country c within a $\pm$2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally include time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

the same discipline and time period. We estimate regressions based on Equation (3) in which the dependent variable is the inverse hyperbolic sine of the number of highly-cited publications in country c and time t (i.e., the $\pm$2-year window around year t).

We find a strong, positive, and significant relationship between institutional quality and the production of frontier research independently of whether we consider publications in the top 10% (Panel A), top 5% (Panel B), top 3% (Panel C), or top 1% (Panel D) of the citation distribution as frontier research (Table 5). The coefficient estimate of 1.445 (Panel A, column 4) for the top 10% research implies that a 0.1 change in the V-Dem Liberal Democracy Index is associated with a 16% change in the number of publications that end up in the top 10% of the discipline-level citation distribution.

Overall, these results suggest that stronger institutions shape not only the volume of scientific publications a country produces, but also its output of highly-cited research, the segment of scientific activity that disproportionately drives progress in science and technology (e.g., Lotka 1926; Merton 1957; Iaria et al. 2018).

Nobel Prize-Winning Discoveries

While highly-cited publications measure the upper tail of scientific output, they may still not capture the most transformative scientific discoveries. To assess whether political institutions also influence truly groundbreaking discoveries, we study effects on Nobel Prize-winning research. These discoveries capture the most significant scientific achievements made in physics, chemistry, and medicine throughout the 20$^{\text{th}}$ century and beyond. Notably, Nobel Prizes are awarded to scientists even when they work in highly repressive

Table 5: Political Institutions and Highly-Cited Publications

	Dep. Var.: Asinh Highly-Cited Publications			
	(1)	(2)	(3)	(4)
Panel A: Top 10% Research				
Institutional Quality	6.960***	5.710***	1.694***	1.445***
	(0.288)	(0.376)	(0.406)	(0.376)
Observations	3,774	3,774	3,774	3,774
R^2	0.688	0.756	0.903	0.913
Dependent Variable Mean	3.00	3.00	3.00	3.00
Panel B: Top 5% Research				
Institutional Quality	6.462***	5.400***	1.635***	1.359***
	(0.294)	(0.379)	(0.428)	(0.392)
Observations	3,774	3,774	3,774	3,774
R^2	0.665	0.726	0.881	0.895
Dependent Variable Mean	2.48	2.48	2.48	2.48
Panel C: Top 3% Research				
Institutional Quality	6.055***	5.130***	1.566***	1.269***
	(0.301)	(0.380)	(0.438)	(0.396)
Observations	3,774	3,774	3,774	3,774
R^2	0.646	0.701	0.862	0.881
Dependent Variable Mean	2.14	2.14	2.14	2.14
Panel D: Top 1% Research				
Institutional Quality	5.094***	4.444***	1.414***	1.085***
	(0.311)	(0.377)	(0.436)	(0.386)
Observations	3,774	3,774	3,774	3,774
R^2	0.593	0.638	0.818	0.850
Dependent Variable Mean	1.51	1.51	1.51	1.51
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of variants of Equation (3), where the dependent variable is the inverse hyperbolic sine of the number of highly-cited publications in country c within a $\pm$2-year window around year t. Panel A presents results for publications in the top 10%, Panel B for the top 5%, Panel C for the top 3%, and Panel D for the top 1% of the discipline and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally include time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

political environments and often decades after completing their research. For instance, several scientists from Imperial Russia and the Soviet Union received Nobel Prizes. Ivan Petrovich Pavlov, for example, received the 1904 Nobel Prize in Physiology/Medicine for his experimental research using dogs to study "the physiology of digestion" (Nobel Foundation, 2025a). We estimate the following regressions:

$$\text{Nobel Discoveries}_{ct} = \phi \cdot \text{Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \quad (4)$$

where $\text{Nobel Discoveries}_{ct}$ is the weighted number of Nobel Prize-winning discoveries produced in country c in time t (i.e., the $\pm$2-year window around year t). As described in Section 2.4, we weight Nobel Prize-winning discoveries by the number of prize winners and the corresponding country affiliations that led to the award. The measure captures the time and place where Nobel Prize-winning research was produced and not the year and location of the scientist at the time of the award. As before, we control for time fixed effects. While the total number of Nobel Prizes awarded per year does not increase over time, there are certain time trends that may affect the amount of Nobel-winning research over time; e.g., the interruption of Nobel Prize awards during World War I and World War II. As Nobel Prizes are awarded with a delay, these interruptions mechanically reduced the amount of Nobel Prize-winning research in pre-WWI and pre-WWII years. We also control for country fixed effects, which account for the fact that some countries systematically produce more Nobel Prize-winning research.

Table 6: Political Institutions and Nobel Prize-Winning Discoveries

	Dep. Var.: Nobel Discoveries			
	(1)	(2)	(3)	(4)
Institutional Quality	0.377**	0.575**	0.157*	0.177*
	(0.170)	(0.252)	(0.0864)	(0.0931)
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes
Observations	3,774	3,774	3,774	3,774
R^2	0.055	0.089	0.583	0.585
Dependent Variable Mean	0.08	0.08	0.08	0.08

Notes: The table reports estimates of Equation (4), where the dependent variable is the weighted number of Nobel Prize-winning discoveries in country c within a $\pm$2-year window around year t. The discoveries are weighted by the number of prize winners and the country affiliations reported in the Nobel-winning publications (see Section 2.4 for details). The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

We find that better political institutions are associated with a higher rate of Nobel Prize-winning discoveries (Table 6). An improvement of institutional quality by 0.1 points increases these discoveries by 0.018 or 21% (Table 6, column 4, p-value 0.059).

These findings highlight that political institutions shape the production of breakthrough discoveries even at the very top of the quality distribution. Moreover, to the extent that the Nobel Committee ultimately recognizes most major scientific advances, the results imply that weaker institutions depress the absolute rate of path-breaking discoveries rather than merely reflecting unobserved research activity in authoritarian regimes. We return to this issue in the last part of the paper.

4.2 Institutions and Scientific Output: Effect of Sudden Changes in Political Institutions

The previous results establish a strong relationship between a country's political institutions and both the quantity and quality of scientific output. To investigate a causal link between political institutions and scientific output, we examine the effect of sudden and major changes in institutional quality that occurred in different countries during the 20$^{\text{th}}$ and 21$^{\text{st}}$ centuries. The global coverage of our data allows us to systematically identify major institutional shocks, defined as all episodes in which a country's V-Dem Liberal Democracy Index changed by at least 0.2 points (either sudden improvements or sudden deteriorations) within a five-year period. A shift of this magnitude corresponds to a substantial change in institutional quality, comparable in size to the transformations observed during major democratization or autocratization episodes.[19]

Since 1900, there have been 77 major improvements, including the democratizations after the fall of the Iron Curtain and Tunisia's transition following the Arab Spring. Over the same time period, we observe 31 major deteriorations, including Franco's fascist dictatorship in Spain, Hitler's Nazi dictatorship in Germany, Batista's military coup that plunged Cuba into dictatorship and paved the way for decades of communist rule, and Orban's authoritarian turn in Hungary. These shocks typically arise from discrete political events, such as regime transitions, coups, or constitutional breakdowns.

We investigate the effect of these sudden changes in political institutions by estimating event study regressions with staggered treatments separately for episodes with a major improvement or deterioration of political institutions. For major improvements, we estimate the following regression:

$$\begin{aligned}\text{asinh}(\text{Output})_{ct} = &\sum_{\tau=-4}^{5} \beta_\tau \text{Positive Institutional Shock}_c \times \mathbf{1}[t(c)=\tau] \\ &+ \text{Country FE}_c + \text{Time FE}_t + X_{ct}\beta + \varepsilon_{ct}\end{aligned} \tag{5}$$

For major deteriorations, we estimate the equivalent regression:

[19]The results are robust to using alternative cutoffs to define major institutional changes (Appendix Figure C.3 and C.4)

$$\begin{aligned}\text{asinh(Output)}_{ct} = &\sum_{\tau=-4}^{5} \beta_\tau \text{Negative Institutional Shock}_c \times \mathbf{1}[t(c) = \tau] \\ &+ \text{Country FE}_c + \text{Time FE}_t + X_{ct}\beta + \varepsilon_{ct}\end{aligned} \tag{6}$$

As before, $\text{asinh(Output)}_{ct}$ is the inverse hyperbolic sine of publications (or highly-cited publications) in country c and time t (i.e., the ± 2-year window around year t). Positive Institutional Shock$_c$ and Negative Institutional Shock$_c$ are indicators that equal one if a country experienced a large improvement or deterioration in political institutions over a five-year period (i.e., an increase or decrease in the LDI between $t-1$ and t of at least 0.2 points). We interact this variable with indicators for the time relative to the treatment to analyze the dynamics of major changes in institutional quality. The regressions also control for country and time fixed effects as well as the inverse hyperbolic sine of the population. As before, we cluster the standard errors at the country level.

The event-study design mitigates concerns that secular trends in scientific activity may be correlated with changes in institutional quality. Moreover, the magnitude and abruptness of the institutional shocks make it implausible that they are themselves driven by changes in scientific output, thereby substantially reducing concerns about reverse causality. The underlying identifying assumption is that, absent these sudden institutional changes, scientific output in affected countries would have followed the same trend as in countries that did not experience sharp institutional changes.

To isolate clean event-study episodes and avoid contamination from overlapping shocks, we exclude country-periods once the Liberal Democracy Index (LDI) in a treated country reverts to its pre-shock level. For example, Germany experienced a sharp improvement in institutional quality with the transition from Imperial Germany to the Weimar Republic after World War I, with the LDI improving from 0.22 to 0.46 over a five-year period. This positive shock was more than reversed in 1933 with the rise of the Nazi regime (the LDI declined from 0.50 to 0.02). In such cases, we define the initial improvement as the treatment event (e.g., the positive "Weimar" shock in Germany) and construct both leads (pre-periods: 1900, 1905, 1910, 1915) and lags (post-periods: 1920, 1925, 1930) relative to that event. However, we exclude all observations after the reversal (in this case, all post-1933 observations for Germany) to avoid contamination from the subsequent negative shock.[20] A small number of countries experienced multiple large institutional shocks in the same direction. In these cases, we retain only the first major shock in each direction.[21]

We first report results using the canonical TWFE difference-in-difference estimator. To address concerns about the staggered treatment timing and non-homogeneous treatment effects, we also estimate the same specification using four alternative estimators as proposed

[20]Note, we separately analyze the negative shock of the Nazi regime by estimating Equation (6), treating the Weimar period (1920, 1925, 1930) as the pre-period and the Nazi era (1935, 1940, 1945) as the post-shock treatment period.

[21]For example, Peru experienced two major institutional deteriorations: the 1968–1980 military dictatorship and the authoritarian period under Fujimori in the 1990s. As we focus on the first negative shock in each country, we estimate effects for the 1968–1980 dictatorship in Peru.

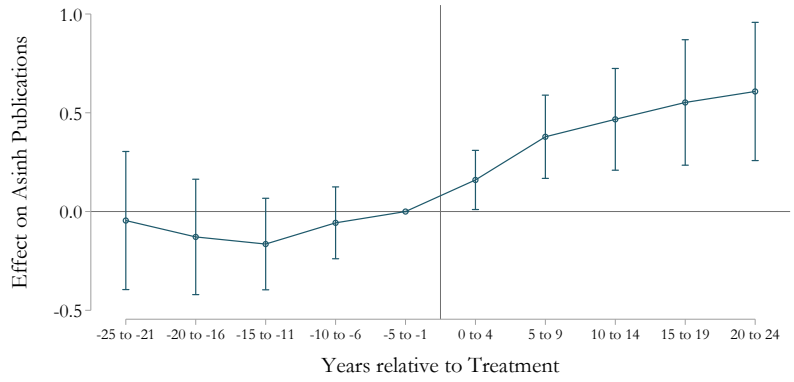
by de Chaisemartin and d'Haultfoeuille (2020), Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Borusyak et al. (2024).

Sudden Improvements of Political Institutions

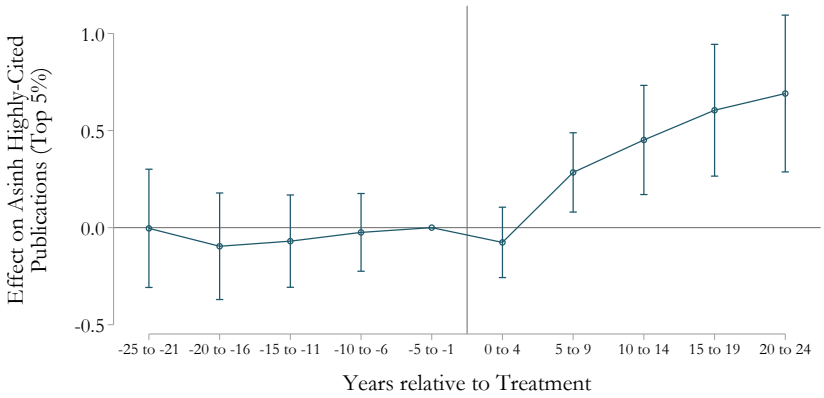
We begin by investigating sudden *improvements* in political institutions and find a large and statistically significant increase in the number of publications following such improvements (Figure 5, Panel a). The effect grows gradually over time and persists for at least 25 years after the shock. The point estimates imply that, 20–24 years after a positive institutional shock, scientific output increases by approximately 84%.

Figure 5: Event Study Estimates Following Sudden Improvement in Institutions

(a) TWFE Event Study: Publications



(b) TWFE Event Study: Highly-Cited Publications (Top 5%)



Notes: Panels (a) and (b) plot point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5) where the dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited publications (panel b) in country c within a ± 2-year window around year t.

We also find a large and significant increase in frontier research after a sudden improvement in institutional quality (Panel b). The point estimate indicates that 20–24 years after a positive institutional shock, the country produces around 100% more highly-cited scientific publications. These results are robust to applying alternative definitions of highly-cited

publications, whether based on the top 10%, top 3%, or top 1% of the discipline and year citation distribution (Appendix Figure C.1).

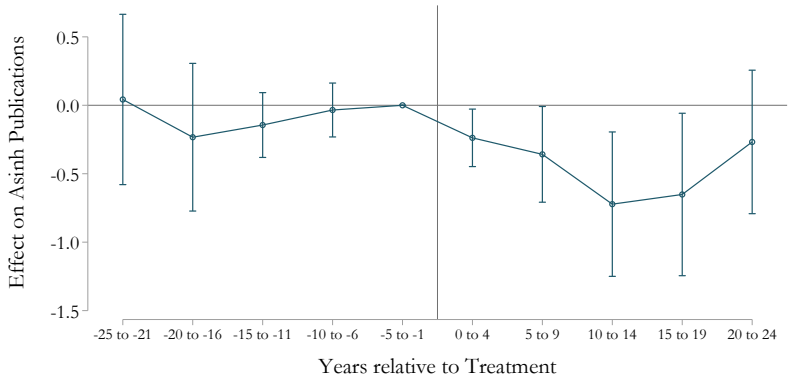
The estimates of the pre-period coefficients indicate that countries subsequently experiencing large improvements in institutional quality exhibit similar trends in the 25 years preceding the shock. This pattern suggests that these countries did not experience differential shocks prior to the institutional change, supporting the identifying assumption.

Sudden Deteriorations of Political Institutions

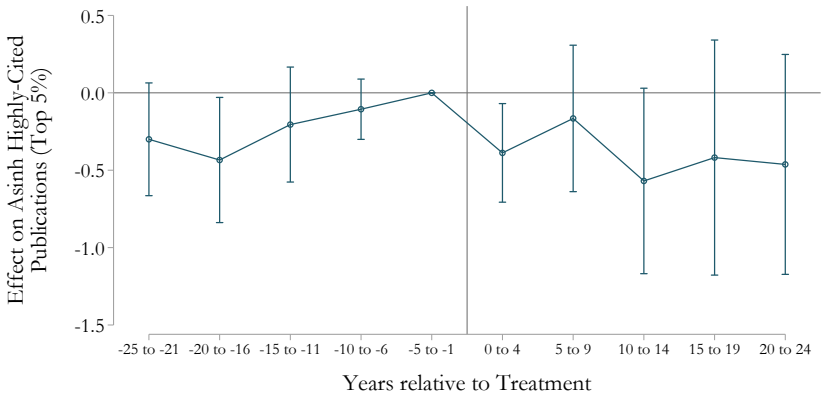
Next, we investigate sudden *deteriorations* in the quality of political institutions. After a major deterioration, the number of publications gradually declines, with the largest decline around 10–14 years after the shock. The negative effect on publications persists for about 20 years (Figure 6, Panel a). The corresponding estimates exhibit larger standard errors, probably because negative institutional shocks occur less frequently than positive ones.

Figure 6: Event Study Estimates Following Sudden Deterioration in Institutions

(a) TWFE Event Study: Publications



(b) TWFE Event Study: Highly-Cited Publications (Top 5%)



Notes: Panels (a) and (b) plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (6) where the dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited publications (panel b) in country c within a $\pm$2-year window around year t.

We also find a similar decline for the number of frontier publications (Panel b). For this outcome, only the first of the individual post-period coefficients is significant at the 5%-level, but all point estimates for the post-event period are negative and indicate a decline in frontier research. These results are similar if we apply alternative definitions of highly-cited publications, whether based on the top 10%, top 3%, or top 1% of the discipline and year citation distribution (Appendix Figure C.2).

The pre-treatment coefficient estimates indicate that treated and control countries exhibit comparable trends in the 25 years preceding a sudden deterioration in institutional quality, providing further support for the identifying assumption. For highly-cited publications (Panel B), we even observe a modest upward trend in the pre-period. This indicates that, if anything, frontier scientific output was improving before countries experienced a sudden institutional deterioration, suggesting that we estimate a lower bound.

Overall, the point estimates of positive and negative shocks are of comparable magnitude and opposing signs. This indicates that improvements and deteriorations in institutional quality have symmetric but opposing effects on scientific output.

Robustness

The findings are robust to using alternative estimators (de Chaisemartin and d'Haultfoeuille 2020, Callaway and Sant'Anna 2021, Sun and Abraham 2021, and Borusyak et al. 2024) proposed for event-studies with staggered treatment timing and dynamic treatment effects (Figures 7 and 8).

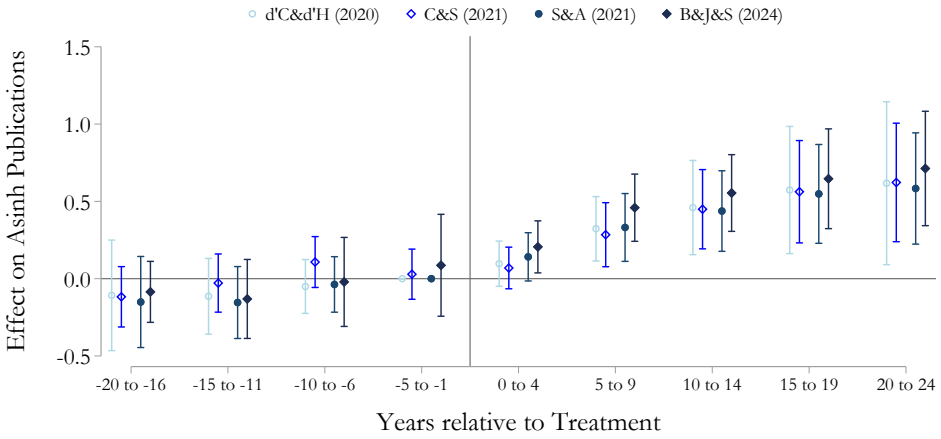
We also show that the results are robust to using alternative thresholds to define sudden institutional changes. In the main analysis, we classify major institutional shocks as episodes in which a country's V-Dem Liberal Democracy Index (LDI) changes by at least 0.2 points within a five-year window. The results are similar if we instead define major shocks as episodes with LDI changes of at least 0.15 points (Appendix Figure C.3) or at least 0.25 points (Appendix Figure C.4).

In an additional set of robustness checks, we show that concurrent episodes of conflict do not drive our results. Institutional change may, in some instances, be correlated with internal or external conflict that could directly influence scientific production. For example, the transition from Weimar Germany to the Nazi regime in 1933 was followed by the outbreak of WWII six years after the sharp institutional shock. To address this concern, we re-estimate the event study regressions while controlling for episodes of conflict. We measure exposure to conflict using data from the Correlates of War Project (Sarkees and Wayman, 2010).

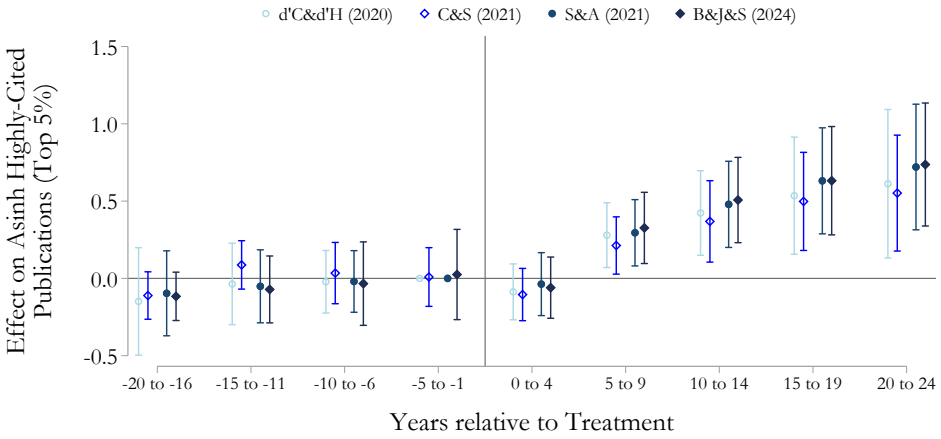
The Correlates of War Project reports all interstate wars, such as World War I and World War II, together with the participating countries, the years of involvement, and the associated number of battle deaths. Using these data, we construct a country-year measure of exposure to interstate conflict. For each war, we first compute average annual battle deaths for each participating country and assign this value to all years in which the country

Figure 7: Staggered Event Study Estimators – Sudden Improvements

(a) Alternative Estimators: Publications



(b) Alternative Estimators: Highly-Cited Publications (Top 5%)

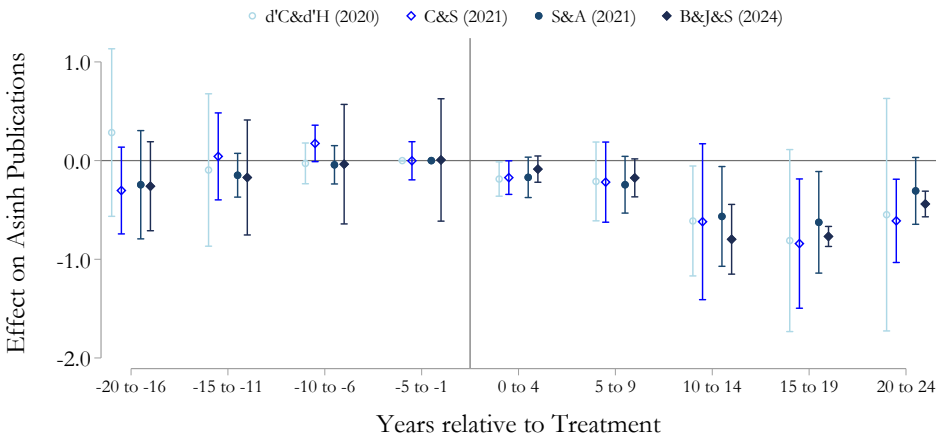


Notes: The figure plots the point estimates and 95% confidence intervals of the staggered difference-in-differences estimators developed by de Chaisemartin and d'Haultfoeuille (2020) [d'C&d'H (2020)], Callaway and Sant'Anna (2021) [C&S (2021)], Sun and Abraham (2021) [S&A (2021)], and Borusyak et al. (2024) [B&J&S (2024)] for the event study in Equation (5) following a sudden improvement in institutional quality. The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited publications (panel b) in country c within a ±2-year window around year t.

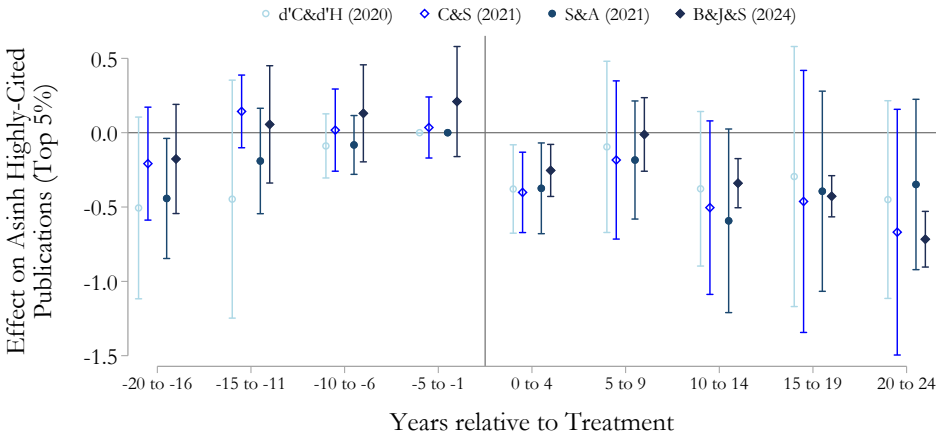
was at war, and zero otherwise. Because our publication outcomes are measured in a ±2-year window around year t, we count the number of years a country was exposed within the ±2-year window and scale the resulting measure by population size. For example, Germany was engaged in World War II from 1939 to 1945 and suffered approximately 3.5 million battle deaths, or about 500,000 per year. For the 1940 observation (which considers publications between 1938 and 1942), the war exposure variable equals [4 (years) × 500,000 (battle deaths)]/68,829.71 (population in thousands) = 29.06. For the 1945 observation, it equals [3 (years)×500,000 (battle deaths)]/65,678.29 (population in thousands) = 22.84. In contrast, the United States entered World War II only in 1941 and incurred 405,400 battle deaths overall, or 81,080 per year. Accordingly, the war exposure variable

Figure 8: Staggered Event Study Estimators – Sudden Deteriorations

(a) Alternative Estimators: Publications



(b) Alternative Estimators: Highly-Cited Publications (Top 5%)



Notes: The figure plots the point estimates and 95% confidence intervals of the staggered difference-in-differences estimators developed by de Chaisemartin and d'Haultfoeuille (2020) [d'C&d'H (2020)], Callaway and Sant'Anna (2021) [C&S (2021)], Sun and Abraham (2021) [S&A (2021)], and Borusyak et al. (2024) [B&J&S (2024)] for the event study in Equation (6) following a sudden deterioration in institutional quality. The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited publications (panel b) in country c within a $\pm$2-year window around year t.

is much lower: 1.23 in 1940 and 1.72 in 1945.[22] Using this newly constructed control variable, we re-estimate the event-study regressions and control for exposure to interstate wars. The results closely mirror our baseline findings (Appendix Figure C.5).

Sharp changes in political institutions may also coincide with intrastate or civil wars. To address this possibility, we construct a measure of exposure to intrastate conflict. As before, we use the Correlates of War database to identify the participating countries, years of participation, and the associated number of battle deaths. We then re-estimate

[22]For 1940, which considers publications between 1938 and 1942 (and thus war exposure in 1941 and 1942), the calculation is [2 (years) $\times$ 81,080 (battle deaths)]/131,669 (population in thousands) = 1.23. For 1945, it is [3 (years) $\times$ 81,080 (battle deaths)]/141,183 (population in thousands) = 1.72.

the event-study regressions and control for this measure of intrastate war exposure. The results remain very similar to our baseline findings (Appendix Figure C.6).

Overall, these results indicate that conflict occurring around periods of sharp institutional change is not the primary driver of the observed shifts in scientific output. This suggests that our findings are not simply capturing the disruptive consequences of conflict but instead reflect an independent role of political institutions in shaping the production of scientific research.

Finally, we examine whether the results are disproportionately driven by institutional changes on any particular continent. For instance, one might be concerned that the effects are largely attributable to the collapse of Communism and subsequent democratization episodes in Europe. For these tests, we re-estimate the event studies while sequentially excluding data from one continent at a time. The findings show that the results are not driven by institutional changes in any single region (Appendix Figures C.7 and C.8).

5 Decomposing Changes in Scientific Production: Accounting for Size Differences

In the preceding part, we have documented that scientists in countries with higher-quality political institutions produce more scientific publications and frontier research. Better institutions may foster scientific progress either by expanding the number of researchers or by increasing the productivity per researcher. To distinguish between these mechanisms, we test whether the relationship between political institutions and research output persists once we account for the growth in the number of researchers documented in the first part of the paper. We estimate modified versions of Equation (3) in which we additionally control for the inverse hyperbolic sine of the number of unique authors in country c in a $\pm$2-year window around year t.[23]

Controlling for the size of a country's academic sector, the estimated coefficients on institutional quality shrink by a factor of 5–10 (Table 7). This pattern suggests that a substantial share of the association between institutions and total publications operates through changes in the size of the scientific workforce. Yet, the coefficients on institutional quality remain positive, statistically significant, and economically meaningful, implying that better institutions also raise output per researcher.

We next report regression estimates where the dependent variable is the number of highly-cited publications (Table 8). Once we account for the number of authors, the estimated coefficients on institutional quality fall relative to the baseline, but only by about

[23]Note that we control for the number of unique authors in the OpenAlex data instead of the number of academics from the World of Academia Database (Part 1) because the World of Academia data only captures the period 1900 to 1969 and because these data focus on university academics, whereas OpenAlex also captures publications by more junior researchers, scientists in research institutes, and researchers working in the private sector.

Table 7: Political Institutions and Publications – Controlling for the Number of Authors

	Dep. Var.: Asinh Publications							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Institutional Quality	7.762***	0.745***	5.961***	0.554***	1.317***	0.294**	1.217***	0.244**
	(0.279)	(0.0920)	(0.331)	(0.0784)	(0.312)	(0.117)	(0.305)	(0.113)
Asinh Authors		0.850***		0.914***		0.849***		0.844***
		(0.00705)		(0.00964)		(0.0151)		(0.0150)
Asinh Population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE			Yes	Yes	Yes	Yes	Yes	Yes
Country FE					Yes	Yes	Yes	Yes
GDP per capita							Yes	Yes
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.726	0.986	0.823	0.988	0.946	0.991	0.947	0.991
Dependent Variable Mean	5.30	5.30	5.30	5.30	5.30	5.30	5.30	5.30

Notes: The table reports estimates of variants of Equation (3), where the dependent variable is the inverse hyperbolic sine of the number of publications in country c within a ±2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. Even columns add a control for the inverse hyperbolic sine of the number of unique authors in country c within a ±2-year window around year t in the *OpenAlex* data. All regressions control for the inverse hyperbolic sine of population. In columns (3) and (4), we additionally control for time fixed effects; in columns (5) and (6), we add country fixed effects; and in columns (7) and (8), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

20–40% (column 4). In contrast to the results on overall publication counts, these findings suggest that the higher production of frontier research in countries with stronger political institutions is not primarily driven by increases in the size of the scientific workforce. Instead, researchers in these countries are more productive in generating highly-cited publications.

Lastly, we estimate regressions where the dependent variable is the weighted number of Nobel Prize-winning discoveries. The estimated coefficients on institutional quality are relatively similar in magnitude to those from specifications that do not control for the size of the academic workforce (Table 9). This suggests that effects on Nobel Prize-winning discoveries are mostly driven by increases in the productivity per researcher and not by size effects.

Taken together, these results show that the expansion of the scientific workforce accounts for much of the link between institutions and the quantity of publications, but not for frontier research and breakthrough discoveries. This suggests that while institutions shape the quantity of scientific output partly through their effect on scale, they are particularly important for fostering the quality of research at the frontier, where freedom to explore novel ideas and free experimentation may be particularly important (e.g. Polanyi, 1962).

6 Political Institutions and Idea Diversity

In the fourth part of the paper, we examine whether political institutions shape the breadth of research. The preceding results show that institutions influence both the quantity and quality of scientific output. Yet their impact may extend beyond the overall level of

Table 8: Political Institutions and Highly-Cited Publications – Controlling for the Number of Authors

	Dep. Var.: Asinh Highly-Cited Publications							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Top 10% Research								
Institutional Quality	6.960***	1.931***	5.710***	1.646***	1.694***	1.063***	1.445***	0.865***
	(0.288)	(0.258)	(0.376)	(0.238)	(0.406)	(0.329)	(0.376)	(0.303)
Asinh Authors		0.609***		0.687***		0.524***		0.503***
		(0.0202)		(0.0261)		(0.0384)		(0.0360)
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.688	0.882	0.756	0.891	0.903	0.927	0.913	0.935
Dependent Variable Mean	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
Panel B: Top 5% Research								
Institutional Quality	6.462***	2.107***	5.400***	1.813***	1.635***	1.118***	1.359***	0.891***
	(0.294)	(0.288)	(0.379)	(0.267)	(0.428)	(0.373)	(0.392)	(0.341)
Asinh Authors		0.528***		0.606***		0.429***		0.406***
		(0.0222)		(0.0281)		(0.0408)		(0.0379)
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.665	0.836	0.726	0.850	0.881	0.900	0.895	0.912
Dependent Variable Mean	2.48	2.48	2.48	2.48	2.48	2.48	2.48	2.48
Panel C: Top 3% Research								
Institutional Quality	6.055***	2.190***	5.130***	1.902***	1.566***	1.128***	1.269***	0.879**
	(0.301)	(0.306)	(0.380)	(0.286)	(0.438)	(0.394)	(0.396)	(0.358)
Asinh Authors		0.468***		0.545***		0.364***		0.338***
		(0.0231)		(0.0286)		(0.0418)		(0.0385)
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.646	0.800	0.701	0.816	0.862	0.878	0.881	0.895
Dependent Variable Mean	2.14	2.14	2.14	2.14	2.14	2.14	2.14	2.14
Panel D: Top 1% Research								
Institutional Quality	5.094***	2.241***	4.444***	1.976***	1.414***	1.109***	1.085***	0.827**
	(0.311)	(0.329)	(0.377)	(0.310)	(0.436)	(0.411)	(0.386)	(0.366)
Asinh Authors		0.346***		0.417***		0.253***		0.224***
		(0.0231)		(0.0284)		(0.0416)		(0.0375)
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.593	0.711	0.638	0.732	0.818	0.829	0.850	0.859
Dependent Variable Mean	1.51	1.51	1.51	1.51	1.51	1.51	1.51	1.51
Asinh Population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE			Yes	Yes	Yes	Yes	Yes	Yes
Country FE					Yes	Yes	Yes	Yes
GDP per capita							Yes	Yes

Notes: The table reports estimates of variants of Equation (3), where the dependent variable is the inverse hyperbolic sine of the number of highly-cited publications in country c within a $\pm$2-year window around year t. Panel A presents results for publications in the top 10%, Panel B for the top 5%, Panel C for the top 3%, and Panel D for the top 1% of the discipline and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. Even columns add a control for the inverse hyperbolic sine of the number of unique authors in country c within a $\pm$2-year window around year t in the *OpenAlex* data. All regressions control for the inverse hyperbolic sine of population. In columns (3) and (4), we additionally control for time fixed effects; in columns (5) and (6), we add country fixed effects; and in columns (7) and (8), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Political Institutions and Nobel Prize-Winning Discoveries – Controlling for the Number of Authors

	Dep. Var.: Nobel Prize-Winning Discoveries							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Institutional Quality	0.377**	0.451**	0.575**	0.312**	0.157*	0.129*	0.177*	0.148*
	(0.170)	(0.182)	(0.252)	(0.127)	(0.0864)	(0.0743)	(0.0931)	(0.0807)
Asinh Authors		-0.00899**		0.0445*		0.0235*		0.0255*
		(0.00426)		(0.0243)		(0.0130)		(0.0133)
Asinh Population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE			Yes	Yes	Yes	Yes	Yes	Yes
Country FE					Yes	Yes	Yes	Yes
GDP per capita							Yes	Yes
Observations	3,774	3,774	3,774	3,774	3,774	3,774	3,774	3,774
R^2	0.055	0.056	0.089	0.102	0.583	0.584	0.585	0.586
Dependent Variable Mean	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08

Notes: The table reports estimates of variants of Equation (4), where the dependent variable is the weighted number of Nobel Prize-winning discoveries in country c within a ±2-year window around year t. The discoveries are weighted by the number of prize winners and the country affiliations reported in the Nobel-winning publications (see Section 2.4 for details). The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. Even columns add a control for the inverse hyperbolic sine of the number of unique authors in country c within a ±2-year window around year t in the *OpenAlex* data. All regressions control for the inverse hyperbolic sine of population. In columns (3) and (4), we additionally control for time fixed effects; in columns (5) and (6), we add country fixed effects; and in columns (7) and (8), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

research to its distribution across fields. Philosophers and sociologists of science (e.g., Merton, 1942; Popper, 1945, p. 426; Polanyi, 1962) have argued that authoritarian countries produce narrower research portfolios than democracies because democratic regimes allow scientists to freely explore a wide range of ideas without the risk of political interference. Similarly, authoritarian countries may invest public funds in a narrower range of fields that are deemed particularly promising. Both of these mechanisms would generate a more concentrated research portfolio under authoritarian regimes.

To test this hypothesis, we quantify the degree of research specialization at the country-year level. Specifically, we construct concentration ratios based on the distribution of publications across 256 scientific fields in the OpenAlex data. For each country c and the ± 2-year period around t, the concentration ratio is defined as:

$$Concentration\ Ratio_{ct}^{k} = \sum_{f=1}^{k} s_{fct}$$

where s_{fct} denotes the share of publications in field f for country c in year t. The parameter k determines how many of the largest fields are included when calculating the concentration ratio. For example, $Concentration\ Ratio_{ct}^{25}$ measures the combined share of publications in the 25 largest fields of a given country-year. Using these concentration ratios, we estimate the following regressions:

$$Concentration\ Ratio_{ct}^{k} = \phi \cdot \text{Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \quad (7)$$

where $Concentration\ Ratio_{ct}^{k}$ is the concentration ratio of country c at time t. The other variables are defined as before. Our baseline choice is $k = 25$, so that the concentration

ratio reflects the combined share of the 25 largest fields out of 256 scientific fields in the OpenAlex data. We calculate the concentration ratio separately for all publications and for highly-cited publications based on different cutoffs (top 10%, top 5%, top 3%, and top 1%). Note that concentration ratios can only be computed for country-year observations with at least one (highly-cited) publication. Consequently, countries and periods without any (highly-cited) publications are excluded from these regressions.[24]

We first examine the concentration of all publications, irrespective of quality. The coefficients on institutional quality are negative, but they decline in magnitude and become statistically insignificant once we control for country fixed effects (Table 10, Panel A). These estimates indicate that changes in political institutions do not systematically alter the concentration of overall scientific activity.

Next, we examine the concentration of highly-cited research. In contrast to the results on the number of all publications, the estimates for institutional quality are negative and statistically significant for quality cutoffs above the top 5% (Table 10, Panels B–E). For instance, a 0.1-point increase in the Liberal Democracy Index is associated with a 1.3 percentage-point decline in the concentration ratio of top 1% research (Panel E, column 4).

These results show that higher-quality institutions are associated with lower concentration ratios of frontier research and thus a broader distribution of top scientific output. These results indicate that, while autocracies can generate frontier research even at the top 1% level, such output is concentrated in a narrower set of fields than in democracies. To our knowledge, this is the first systematic empirical evidence in support of the hypothesis put forward by philosophers and sociologists of science.

Robustness

The findings are robust to alternative definitions of concentration ratios and to imposing minimum publication thresholds. First, we vary the cardinality k when computing the ratios and find nearly identical results when using the 20 or 50 largest fields (Appendix Table D.1 and Table D.2). Second, we address the concern that higher concentration ratios in authoritarian regimes may arise mechanically because such regimes produce less (frontier) research overall (as documented in Section 4). To investigate this issue, we restrict the sample to country-time observations with at least 250, 500, 1,000, or 2,000 publications in a $\pm$ 2-year window. The results remain highly significant and are somewhat larger in absolute magnitude, even when we limit the analysis to observations with at least 2,000 publications (Appendix Tables D.3 to D.6).

[24]By construction, country-year observations with very few publications yield higher concentration ratios. For example, if researchers in a country only produced one highly-cited publication over a five-year period, the concentration ratio of highly-cited publications would be 1. To address this concern, we demonstrate that our results are robust to imposing a minimum number of publications per country in a $\pm$2-year period around year t (Appendix Tables D.3 to D.6).

Table 10: Political Institutions and Concentration of Ideas

	Dep. Var.: Concentration Ratio[25]			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.384***	-0.296***	-0.013	-0.022
	(0.020)	(0.023)	(0.029)	(0.029)
Observations	3,288	3,288	3,288	3,288
R^2	0.544	0.656	0.836	0.840
Dependent Variable Mean	0.72	0.72	0.72	0.72
Panel B: Top 10% Research				
Institutional Quality	-0.410***	-0.351***	-0.044	-0.044
	(0.019)	(0.024)	(0.031)	(0.031)
Observations	2,625	2,625	2,625	2,625
R^2	0.496	0.594	0.819	0.819
Dependent Variable Mean	0.80	0.80	0.80	0.80
Panel C: Top 5% Research				
Institutional Quality	-0.412***	-0.365***	-0.070**	-0.068**
	(0.020)	(0.024)	(0.034)	(0.034)
Observations	2,401	2,401	2,401	2,401
R^2	0.489	0.585	0.815	0.816
Dependent Variable Mean	0.82	0.82	0.82	0.82
Panel D: Top 3% Research				
Institutional Quality	-0.414***	-0.373***	-0.097***	-0.095***
	(0.020)	(0.025)	(0.034)	(0.034)
Observations	2,252	2,252	2,252	2,252
R^2	0.485	0.575	0.811	0.814
Dependent Variable Mean	0.84	0.84	0.84	0.84
Panel E: Top 1% Research				
Institutional Quality	-0.395***	-0.370***	-0.131***	-0.130***
	(0.021)	(0.024)	(0.033)	(0.033)
Observations	1,916	1,916	1,916	1,916
R^2	0.481	0.550	0.790	0.804
Dependent Variable Mean	0.88	0.88	0.88	0.88
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of Equation (7), where the dependent variable is the concentration ratio (cardinality $k = 25$) of publications or highly-cited publications in country c within a $\pm$2-year window around year t. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

7 Path-Breaking Technological Discoveries

The previous results establish a strong link between political institutions and scientific advancements. However, not all breakthrough discoveries are published in academic papers or books. Certain research, particularly with military or strategic applications, may be classified or concealed for political reasons. Authoritarian regimes, in particular, may disproportionately divert scientific activity away from publication-focused science toward secretive research.

To study whether political institutions affect the creation of new ideas more broadly, we use data on path-breaking technological discoveries in the 20$^{\text{th}}$ century and beyond from Asirvatham et al. (2026). After dropping duplicate observations, the data contain 20,785 discoveries. Importantly, the data also cover path-breaking discoveries that were not published in academic journals or books. For example, military discoveries such as the first atomic warhead (United States, 1945) or the first ballistic rocket (the so-called V2 rocket - Nazi Germany). The data also cover government-led discoveries such as the Soviet space program (Soviet Union), which delivered both the first satellite (1957) and human (1961) into space, as well as commercial discoveries such as the first Microsoft operating system MS-DOS (United States, 1981) or ByteDance's TikTok app (China, 2016). Furthermore, the data also contain path-breaking discoveries that were developed in universities; for example, the discovery of penicillin (United Kingdom 1928), the first heart transplant (South Africa, 1967), or CRIPR-Cas9 gene editing (United States, 2012). Finally, the data cover discoveries by independent inventors such as the invention of the aeroplane by the Wright Brothers (United States, 1907).

We study the relationship between political institutions and breakthrough technological discoveries by estimating the following regression:

$$\text{asinh}(\text{Discoveries})_{ct} = \phi \cdot \text{ Institutional Quality}_{ct} + X_{ct}\beta + \text{Country FE}_c + \text{Time FE}_t + \epsilon_{ct} \qquad (8)$$

where the dependent variable, $\text{asinh}(\text{Discoveries})_{ct}$, is the inverse hyperbolic sine of the number of discoveries in country c and time t. As above, we measure discoveries in a $\pm$2-year window around year t. The cross-country discovery distribution has a long right tail. For example, in the $\pm$2-year window around 1995, U.S. scientists made 1,182 discoveries. We therefore use the inverse hyperbolic sine transformation of the number of discoveries.

We find that countries with better political institutions produce more technological discoveries (Table 11). In the specification that includes country and year fixed effects, as well as a control for GDP, the estimated coefficient is 0.343. The estimates imply that a 0.1-point increase in the V-Dem Liberal Democracy Index, roughly the institutional gap between Italy and Denmark in 2020, is associated with an increase in technological discoveries by around 3.5 percent.

Table 11: Political Institutions and Technological Breakthroughs

	Dep. Var.: Asinh Discoveries			
	(1)	(2)	(3)	(4)
Institutional Quality	1.814***	2.389***	0.359**	0.343**
	(0.303)	(0.356)	(0.154)	(0.150)
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes
Observations	3,774	3,774	3,774	3,774
R^2	0.314	0.410	0.864	0.865
Dependent Variable Mean	0.43	0.43	0.43	0.43

Notes: The table reports estimates of Equation (8), where the dependent variable is the inverse hyperbolic sine of the number of technological breakthroughs in country c within a $\pm$2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

The Asirvatham et al. (2026) data assign each discovery one of five institutional types: universities, government, military, corporate, or independent.[25] We explore whether political institutions have a differential impact on discoveries by using the number of discoveries in each of the five categories as the dependent variable.

Institutional quality has a large positive association with breakthroughs produced by universities (point estimate: 0.326) and by industry (point estimate: 0.311), and a somewhat smaller but still statistically significant association with government-led discoveries (point estimate 0.226) (Table 12, columns 1, 2, and 4). In contrast, the estimates for military discoveries are statistically indistinguishable from zero with a negative point estimate (-0.093). Finally, the coefficient for independent discoveries is close to zero and statistically insignificant.

These patterns suggest that higher-quality institutions primarily drive breakthrough discoveries in parts of the innovation system that rely on open scientific exchange, secure property rights, and private-sector commercialization (universities and industry), with somewhat smaller effects on government-led discoveries. In contrast, autocratic regimes disproportionately promote discoveries in military technologies, even though the difference to democracies is not significantly different from 0. Finally, political institutions have little

[25]The Asirvatham et al. (2026) data does not report an institutional type for 459 discoveries (corresponding to around 2.2% of all discoveries).

discernible impact on independent inventors, whose serendipitous discoveries may rely less on formal institutional quality and more on idiosyncratic individual initiative.

Table 12: Technological Breakthroughs by Institution Type

	Dep. Var.: Asinh Discoveries (1900–2020) by				
	Universities (1)	Government (2)	Military (3)	Industry (4)	Independent (5)
Institutional Quality	0.326***	0.226*	-0.093	0.311*	0.019
	(0.116)	(0.121)	(0.103)	(0.169)	(0.047)
Asinh Population	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
GDP per capita	Yes	Yes	Yes	Yes	Yes
Observations	3,774	3,774	3,774	3,774	3,774
R^2	0.795	0.727	0.651	0.832	0.705
Dependent Variable Mean	0.21	0.19	0.06	0.28	0.09

Notes: The table reports estimates of Equation (8), where the dependent variable is the inverse hyperbolic sine of the number of technological breakthroughs developed in a particular institutional type in country c within a $\pm$2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population, time fixed effects, country fixed effects, and GDP per capita. Standard errors are clustered at the country level. Significance levels: $***$ p<0.01, $**$ p<0.05, $*$ p<0.1.

Robustness

The results are robust to using the original Asirvatham et al. (2026) data without deduplicating discoveries (Appendix Table E.1). The results are also similar when using discovery counts rather than the inverse hyperbolic sine transformation (Appendix Table E.2). As discovery counts have large outliers, we use winsorized discovery counts. The estimated coefficients based on non-winsorized counts are even larger but not statistically significant at conventional levels.

8 Conclusion

This paper presents the first comprehensive, long-term empirical analysis of how political institutions shape the creation of ideas. Drawing on a novel worldwide dataset on academics, universities, publications, and discoveries from 1900 to the present, we document five key findings. First, we document a strong positive association between political institutions and the size of the academic sector. Second, we demonstrate that researchers in countries with stronger political institutions consistently produce more scientific output, particularly frontier research. Event studies that exploit sudden deteriorations or improvements in institutional quality confirm that these relationships are driven by institutional change rather than by selection or long-run trends. Third, we find that democracies foster frontier

research even after holding the size of the scientific workforce constant, indicating higher productivity per researcher. Fourth, we document that autocracies concentrate their research in a narrower set of fields: while they can achieve high-quality research in selected areas, they fail to sustain the broad exploration of ideas that characterizes democracies. Fifth, we demonstrate that political institutions affect technological innovations, beyond scientific discoveries published in scientific journals and books.

Taken together, our findings demonstrate that democratic institutions and political freedoms provide the foundation for sustained scientific advancement. Although autocratic regimes may at times produce episodes of notable innovation, broad and persistent progress is far more likely when researchers operate in an environment that protects autonomy, encourages critical inquiry, and supports intellectual freedom. As Karl Popper (1945, p. 433) emphasizes:

> "Reason, like science, grows by way of mutual criticism; the only possible way of 'planning' its growth is to develop those institutions that safeguard the freedom of this criticism, that is to say, the freedom of thought."

Our results underscore that science flourishes when insulated from political interference and that societies safeguarding academic independence are best positioned to expand the frontier of knowledge and innovation.

References

Abbott, A. (2016). Turkey Purges Universities after Failed Coup. *Nature*. July 19.

Acemoglu, D., C. G. Aksoy, C. Baysan, C. Molina, and G. Zeki (2024). Misperceptions and Demand for Democracy under Authoritarianism. NBER Working Paper 33018.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review 91*(5), 1369–1401.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a Fundamental Cause of Long-Run Growth. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1A, pp. 385–472. Amsterdam: Elsevier.

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2019). Democracy Does Cause Growth. *Journal of Political Economy 127*(1), 47–100.

Acemoglu, D. and J. A. Robinson (2013). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Currency.

Aghion, P. and P. Howitt (1992). A Model of Growth through Creative Destruction. *Econometrica 60*(2), 323–351.

Ahmed, M., M. A. Khan, A. Attique, M. A. Khan, H. Haddad, and N. M. Al-Ramahi (2024). Democracy's Limited Impact on Innovation: Panel Data Evidence from Developing Countries. *PLOS One 19*(3), e0297915.

Asirvatham, H., E. Mokski, and A. Shleifer (2026). GPT as a Measurement Tool. NBER Working Paper 34834.

Audretsch, D. B., C. Fisch, C. Franzoni, P. P. Momtaz, and S. Vismara (2024). Academic Freedom and Innovation. *PLOS One 19*(6), e0304560.

Bai, Y., R. Jia, and J. Yang (2024). Knowledge Suppression and Resilience under Censorship: Three-Century Book Publications in China. NBER Working Paper 33258.

Becker, S. O., V. Lindenthal, S. Mukand, and F. Waldinger (2024). Scholars at Risk: Professional Networks and High-Skilled Emigration from Nazi Germany. *American Economic Journal: Applied Economics 16*(3), 1–45.

Beraja, M., A. Kao, D. Y. Yang, and N. Yuchtman (2023). AI-tocracy. *The Quarterly Journal of Economics 138*(3), 1349–1402.

Beraja, M., D. Y. Yang, and N. Yuchtman (2023). Data-intensive Innovation and the State: Evidence from AI firms in China. *The Review of Economic Studies 90*(4), 1701–1723.

Biagioli, M. (1993). *Galileo, Courtier: The Practice of Science in the Culture of Absolutism*. Chicago: University of Chicago Press.

Borjas, G. J. and K. B. Doran (2015). Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors. *Review of Economics and Statistics 97*(5), 1104–1117.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *Review of Economic Studies 91*(6), 3253–3285.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics 225*(2), 200–230.

Cantoni, D. and N. Yuchtman (2014). Medieval Universities, Legal Institutions, and the Commercial Revolution. *The Quarterly Journal of Economics 129*(2), 823–887.

Clarivate Analytics and Chinese Academy of Sciences (2024). Research Fronts 2024: Report to Identify 125 Research Fronts. Technical report, Clarivate and Chinese Academy of Sciences.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement 20*(1), 37–46.

Coppedge, M., J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, K. L. Marquardt, J. Medzihorsky, D. Pemstein, L. Fox, L. Gastaldi, J. Pernes, O. Rydén, E. Tzelgov, Y.-t. Wang, and S. Wilson (2024). V-Dem Methodology v14. Varieties of Democracy (V-Dem) Project.

Coppedge, M., J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, K. L. Marquardt, J. Medzihorsky, D. Pemstein, J. Pernes, J. von Römer, N. Stepanova, E. Tzelgov, Y.-t. Wang, and S. Wilson (2019). V-Dem Methodology v9. Varieties of Democracy (V-Dem) Project.

Cox, G. W. and V. Figueroa (2025). The Inquisition and the Decline of Science in Spain. *Explorations in Economic History 98*, 101699.

de Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Deighton, B. (2026). China could be the world's biggest public funder of science. *Nature*.

Dewitte, E., F. Drago, R. Galbiati, and G. Zanella (2024). Science under Inquisition: The Allocation of Talent in Early Modern Europe. CEPR Discussion Paper 17644.

Fernandez, F., V. Chykina, and Y. C. Lin (2024). Science at Risk? Considering the Importance of Academic Freedom for STEM Research Production across 17 OECD Countries. *PLOS One 19*(2), e0298370.

Fleiss, J. L. (1971). Measuring Nominal Scale Agreement among Many Raters. *Psychological Bulletin 76*(5), 378–382.

Funke, M., M. Schularick, and C. Trebesch (2023). Populist Leaders and the Economy. *American Economic Review 113*(12), 3249–3288.

Ganguli, I. (2014). Scientific Brain Drain and Human Capital Formation after the End of the Soviet Union. *International Migration 52*(5), 95–110.

Gao, Y., L. Zang, A. Roth, and P. Wang (2017). Does Democracy Cause Innovation? An Empirical Test of the Popper Hypothesis. *Research Policy 46*(7), 1272–1283.

Gapminder (2024a). GDP per Capita in Constant PPP Dollars - Version 31. `https://www.gapminder.org/data/documentation/gd001/`. Accessed: 2025-08-20.

Gapminder (2024b). Population - Version 8. `https://www.gapminder.org/data/documentation/gd003/`. Accessed: 2025-08-21.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2013). Human Capital and Regional Development. *The Quarterly Journal of Economics 128*(1), 105–164.

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004). Do Institutions Cause Growth? *Journal of Economic Growth 9*(3), 271–303.

Grüttner, M. (2024). *Talar und Hakenkreuz: Die Universitäten im Dritten Reich*. Munich: C.H. Beck.

Hartshorne, E. Y. (1937). *The German Universities and National Socialism*. Cambridge, MA: Harvard University Press.

Hayek, F. A. (1945). The Use of Knowledge in Society. *American Economic Review 35*(4), 519–530.

Hsu, I. C. Y. (1964). The Reorganisation of Higher Education in Communist China, 1949–61. *The China Quarterly 19*, 128–160.

Iaria, A., C. Schwarz, and F. Waldinger (2018). Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science. *The Quarterly Journal of Economics 133*(2), 927–991.

Iaria, A., C. Schwarz, and F. Waldinger (2022). Gender Gaps in Academia: Global Evidence over the Twentieth Century. SSRN Working Paper 4150221.

International Association of Universities (1985). *World List of Universities/Liste Mondiale des Universités 1985–1986*. Basingstoke: Macmillan Publishers LTD.

International Association of Universities/UNESCO (2000). *World List of Universities and Other Institutions of Higher Education*. Basingstoke and Oxford: Maxmillan Publisher LTD.

International Association of Universities/UNESCO (2024). World Higher Education Database (WHED) Portal. `https://whed.net/`. Accessed on 2024-06-17.

Jones, B. F. and B. A. Weinberg (2011). Age Dynamics in Scientific Creativity. *Proceedings of the National Academy of Sciences 108*(47), 18910–18914.

Jones, C. I. (1995). R & D-Based Models of Economic Growth. *Journal of Political Economy 103*(4), 759–784.

Kantor, S. and A. Whalley (2014). Knowledge Spillovers from Research Universities: Evidence from Endowment Value Shocks. *Review of Economics and Statistics 96*(1), 171–188.

Kaya, M. (2018). Turkey's Purge of Critical Academia. *Middle East Report 48*(288), 25–28.

Kozlov, M., J. Tollefson, and D. Garisto (2026). US science after a year of Trump: what has been lost and what remains. *Nature 649*(8098), 812–815.

Ku, H. and T. Mu (2025). The Rise of China and the Global Production of Scientific Knowledge. IZA Discussion Paper 17866.

Li, J., Y. Yin, S. Fortunato, and D. Wang (2019). A Dataset of Publication Records for Nobel Laureates. *Scientific Data 6*(33).

Lindgren, M. (2011). Documentation for GDP per Capita by Purchasing Power Parities for Countries and Territories. `https://www.gapminder.org/documentation/documentation/gapdoc001_v9.pdf`. Accessed: 2025-08-20.

Lotka, A. J. (1926). The Frequency Distribution of Scientific Productivity. *Journal of the Washington Academy of Sciences 16*(12), 317–323.

Maloney, W. F. and F. Valencia Caicedo (2022). Engineering Growth. *Journal of the European Economic Association 20*(4), 1554–1594.

Marshall, M. G. and T. R. Gurr (2020). Polity V Project, Political Regime Characteristics and Transitions, 1800-2018. Center for Systemic Peace. `http://www.systemicpeace.org/inscrdata.html`.

Merton, R. K. (1942). Science and Technology in a Democratic Order. *Journal of Legal and Political Sociology 1*(1), 115–126.

Merton, R. K. (1957). Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review 22*(6), 635–659.

Mitchell, B. (2013). *International Historical Statistics*. London: Palgrave Macmillan.

Mokyr, J. (1992). *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press.

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

Mokyr, J. (2005). The Intellectual Origins of Modern Economic Growth. *The Journal of Economic History 65*(2), 285–351.

Moser, P., A. Voena, and F. Waldinger (2014). German Jewish Émigrés and U.S. Invention. *American Economic Review 104*(10), 3222–3255.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1991). The Allocation of Talent: Implications for Growth. *The Quarterly Journal of Economics 106*(2), 503–530.

Nature (1930). Index Generalis: Annuaire Général des Universités, Grandes Écoles, Académies, Archives, Bibliothèques, Institute Scientifiques, Jardins Botaniques et Zoologiques, Musées, Observatoires, Stés Savantes. *Nature 126*(3169), 126–127.

Nobel Foundation (2025a). Nobel Prize in Physiology or Medicine 1904. `https://www.nobelprize.org/prizes/medicine/1904/summary/`. Accessed: 2025-12-11.

Nobel Foundation (2025b). The Nobel Prize in Chemistry 2020. `https://www.nobelprize.org/prizes/chemistry/2020/summary/`. Accessed: 2025-08-20.

Nobel Foundation (2025c). The Nobel Prize in Physics 1921. `https://www.nobelprize.org/prizes/physics/1921/summary/`. Accessed: 2025-08-20.

Nobel Foundation (2025d). The Nobel Prize in Physics 1958. `https://www.nobelprize.org/prizes/physics/1958/summary/`. Accessed: 2025-08-20.

Nord, M., D. Altman, F. Angiolillo, T. Fernandes, A. Good God, and S. I. Lindberg (2025). *Democracy Report 2025: 25 Years of Autocratization – Democracy Trumped?* Gothenburg: University of Gothenburg: V-Dem Institute.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press.

Özatalay, C. (2020). Purge, Exile, and Resistance: Rethinking the Conflict of the Faculties through the Case of Academics for Peace in Turkey. *European Journal of Turkish Studies. Social Sciences on Contemporary Turkey 30*.

Polanyi, M. (1962). The Republic of Science. *Minerva 1*(1), 54–73.

Popper, K. R. (1945). *The Open Society and Its Enemies*. Princeton, NJ: Princeton University Press.

Price, D. J. d. S. (1963). *Little Science, Big Science*. New York: Columbia University Press.

Priem, J., H. Piwowar, and R. Orr (2022). OpenAlex: A Fully-Open Index of Scholarly Works, Authors, Venues, Institutions, and Concepts. *arXiv Preprint arXiv:2205.01833*.

Rodrik, D., A. Subramanian, and F. Trebbi (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth 9*(2), 131–165.

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy 94*(5), 1002–1037.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy 98*(5), 71–102.

Sarkees, M. R. and F. Wayman (Eds.) (2010). *Resort to War, 1816–2007*. Washington, D.C.: CQ Press.

Schvitz, G., L. Girardin, S. Rüegger, N. B. Weidmann, L.-E. Cederman, and K. S. Gleditsch (2022). Mapping the International System, 1886-2019: The CShapes 2.0 Dataset. *Journal of Conflict Resolution 66*(1), 144–161.

Science News Staff (2025a). Monday Trump Tracker: Columbia University Defunded, GOP Likes Moon Rockets, and $25k Buyouts. *ScienceInsider*. March 10.

Science News Staff (2025b). Trump's Proposed Budget Would Mean 'Disastrous' Cuts to Science: Key Research Budgets Would Shrink by One-Third to One-Half in 2026 Spending Plan. *ScienceInsider*. May 2.

Shepherd, C. and J. Li (2025). How China Came to Dominate the World in Renewable Energy. *The Washington Post*. March 3.

Squicciarini, M. P. and N. Voigtländer (2015). Human Capital and Industrialization: Evidence from the Age of Enlightenment. *The Quarterly Journal of Economics 130*(4), 1825–1883.

Sun, L. and S. Abraham (2021). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics 225*(2), 175–199.

The New York Times (2025). Trump Intends to Cancel All Federal Funds Directed at Harvard. *The New York Times*. May 27.

Valero, A. and J. Van Reenen (2019). The Economic Impact of Universities: Evidence from across the Globe. *Economics of Education Review 68*, 53–67.

Waldinger, F. (2010). Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany. *Journal of Political Economy 118*(4), 787–831.

Waldinger, F. (2012). Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany. *Review of Economic Studies 79*(2), 838–861.

Waldinger, F. (2016). Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge. *Review of Economics and Statistics 98*(5), 811–831.

Weinberg, A. M. (1961). Impact of Large-Scale Science on the United States. *Science 134*(3473), 161–164.

Online Appendix to Dictators, Democracies, and Discoveries

The Appendix presents details on data collection and additional results:

- Appendix A provides details on data.
- Appendix B provides additional results on Section 3.
- Appendix C provides additional results on Section 4.
- Appendix D provides additional results on Section 6.
- Appendix E provides additional results on Section 7

A Additional Details on Data

A.1 Additional Details: Number of Universities

Coding University Types

To identify regular universities and technical universities, we develop a GPT-based algorithm that assigns a specific type (e.g., regular university, technical university, or religious higher-education institution) to each university. Specifically, each university is classified into one of the following 16 categories:

1. University
2. Technical University
3. Religious Education Institutions
4. Agriculture and Veterinary Medicine
5. Medicine & Pharmacy & Dental Institutions
6. Art & Music Institutions
7. Mining & Petrol Institutions
8. Architecture Institutions
9. Teachers Education Institutions
10. Economics and Business and Politics & Government (including Home Economics)
11. Administrative Training and Civil Service Training Institutions
12. Sport Institutions
13. Law School & Notary Institutions
14. Military Schools & Police Academies
15. Other Higher Education Institutions
16. Not a higher education institution (e.g., academy of sciences, Observatoire de Paris, Musée de l'Homme)

Specifically, we use the following GPT prompt:

> *Imagine you are an academic economic historian trying to classify the higher education institutions in different languages into the following categories:*
>
> *1: University*
> *2: Technical University*
> *3: Religious Education Institutions*
> *4: Agriculture and Veterinary Medicine*

5: Medicine & Pharmacy & Dental Institutions
6: Art & Music Institutions
7: Mining & Petrol Institutions
8: Architecture Institutions
9: Teachers Education Institutions
10: Economics and Business and Politics & Government (including Home Economics)
11: Administrative Training and Verwaltungsakademie
12: Sport Institutions
13: Law School & Notary Institutions
14: Military Schools & Police Academies
15: Other Higher Education Institutions
16: Not a higher education institution (e.g., academy of sciences, Observatoire de Paris, Musée de l'Homme):

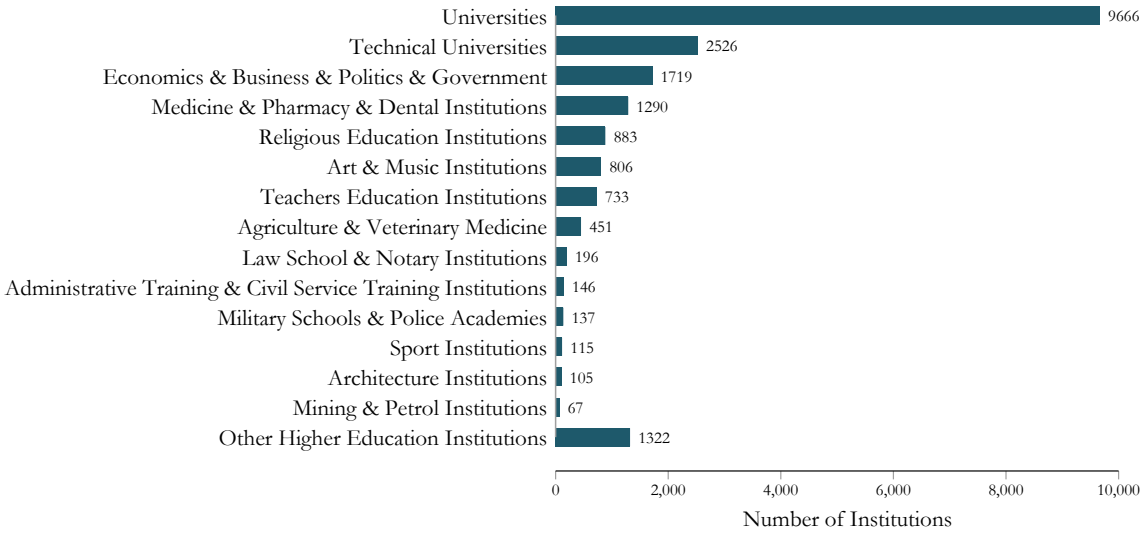
Could you carefully classify the following institutions into the above categories. Only assign unis to 15 if they do not fit into any of the other categories. The category 16 is only for institutions that clearly do not provide higher education (e.g., museums). The following prompts always provides the name of the university followed by a country of the university. Your response should be the number corresponding to the specific university type followed by : and the category name (e.g., 2: Technical University).

We submit each university name to three large language models, GPT-4o, GPT-4.1, and o3-mini, to classify the institution type. Wherever possible, we set the temperature parameter to 0 to minimize randomness in model outputs. After obtaining three predictions (one from each model), we assign the final university type based on the majority classification. In cases where all three models disagree (which occurs in 1,104 of 60,409 instances), we manually code the final university type. The two most common university types are regular universities and technical universities (Figure A.1). For our main results, we define "main universities" as all institutions that are classified as either a regular university or a technical university.

Inter-Model Reliability

Table A.1 reports inter-model reliability for institution-type classifications across the three large language models. Panel A presents pairwise agreement using Cohen's Kappa and Panel B summarizes overall agreement across all three models using Fleiss's Kappa. The high kappa estimates indicate that institution type predictions are consistent across models.

Figure A.1: Number of Higher Education Institutions in 2020 by Type



Notes: The figure shows the total number of higher education institutions by type in 2020.

Table A.1: Inter-Model Reliability of LLM Institution Type Predictions

Panel A: Pairwise Model Comparison (Cohen's Kappa)	Coefficient
GPT-4.1 vs GPT-4o	0.8453
	(0.0016)
GPT-4.1 vs o3-mini	0.8354
	(0.0016)
GPT-4o vs o3-mini	0.8109
	(0.0016)
Panel B: All Models	
Fleiss's Kappa	0.8303
	(0.0014)
Observations	66,859

Notes: Panel A reports pairwise estimates of Cohen's Kappa across all model combinations, based on classifications into 16 institution types. Cohen's Kappa (Cohen, 1960) quantifies inter-rater agreement relative to the agreement expected under statistical independence, taking values in [-1,1], where 1 denotes perfect agreement. Panel B reports Fleiss's Kappa (Fleiss, 1971), which generalizes this framework to multiple raters by measuring the extent of agreement across all three models relative to chance. As with Cohen's Kappa, values are bounded between [-1,1], with higher values indicating greater agreement.

Performance Benchmark: LLM classification versus human assignment of university types

To assess the validity and performance of the LLM-based assignment of university types, we compare the LLMs' predicted majority type classifications to a human classification of a random subset of the 1985 higher education institutions data. We instructed a research

assistant to classify universities into university types. Of the 3,206 higher education institutions, the research assistant coded 1,135 as a main university (regular university or technical university) and 2,071 as another higher education institution. Of the 3,206 higher education institutions, the LLM-based approach classified 1,118 as a main university (regular university or technical university) and 2,018 as other higher education institutions. Table A.2 shows the confusion matrix of the research assistant (RA) and the LLM-based assignments of university type. The F_1 score is 0.8325, indicating good performance of the LLM-based institution assignment.[26]

Table A.2: LLM Majority Predictions and RA Classification

	LLM: Main University	LLM: Other Higher Education	Sum
RA: Main University	967 (85%)	168 (15%)	1,135
RA: Other Higher Education	221 (11%)	1,850 (89%)	2,071
Sum	1,188	2,018	3,206

Notes: The table reports the cross-tabulation of LLM-based classifications and research assistant (RA) classifications of institution type. Entries are counts, with row percentages reported in parentheses. "LLM: University" reports observations classified by the LLMs as universities or technical universities, while "LLM: not University" reports observations classified by the LLMs as a different higher education institution. "RA: University" and "RA: not University" denote human-coded classifications. The sample is restricted to higher education institutions for which institution types were coded by a research assistant.

Imputing the Correct Number of Universities for each Time Period

As outlined in Section 2.1, we combine data from the World of Academia, the World List of Universities, and the World Higher Education Database (WHED) to construct a panel of the number of universities by country over time. A central measurement challenge is that contemporaneous sources may undercount the number of universities, in particular because newly established universities may enter the records only with a delay.

We address this issue by exploiting information on universities' founding dates to reconstruct the stock of institutions in each country at each historical point in time. Specifically, we use founding year data from later cross-sections (1914, 1925, 1938, 1956, and 1969 from World of Academia; 1985 and 2000 from the World List of Universities; and 2020 from WHED) to infer the number of universities that must have existed in a country in earlier years.[27] For a given year t (e.g., 1900), we identify all universities observed in subsequent cross-sections (e.g., 1914 and 1925) whose founding dates precede t, thereby recovering institutions that are missing from the contemporaneous records.

This approach generates multiple measures of the number of universities for each country-year, combining information from contemporaneous and future cross-sections. We take the maximum across these measures as our preferred measure.

To limit potential overcounting, we restrict the backfilling procedure to information from subsequent snapshots observed within 25 years of year t. This addresses the concern

[26] $F_1 = \frac{2 \cdot \text{True Positive}}{2 \cdot \text{True Positive} + \text{False Positive} + \text{False Negative}} = \frac{2 \cdot 967}{2 \cdot 967 + 221 + 168} \approx 0.8325$.

[27] To ensure that we assign universities to the correct historical country in the backfilling procedure, we geocode each university at the city-level and assign countries for earlier periods using yearly shape files of historical country borders (Schvitz et al., 2022).

that some institutions were initially established as specialized schools and only later became regular universities. These changes in university types, could lead to upward-biased university counts if we backfilled information from distant future snapshots.

To illustrate, suppose a country reported 20 universities in $t = 1900$, while the 1914 cross-section lists 30 universities, of which 22 were founded before 1900, and the 1925 cross-section lists 50 universities, of which 21 were founded before 1900. Our procedure would assign 22 universities to 1900, corresponding to the maximum implied count across cross-sections that lie within 25 years of year t.

While the backfilling procedure reduces measurement error, the results do not depend on choosing a particular backfilling window. In fact, results are robust to alternative windows (20 years, 30 years, 50 years, or 120 years) and to only using contemporaneous counts without backfilling (see Appendix Table B.3).

A.2 Additional Details: Technological Discoveries

As described in the main text, we remove duplicate entries from the Asirvatham et al. (2026) data. We identify duplicates based on the semantic similarity of discovery names, measured using sentence embeddings. Sentence embeddings map text strings into a fixed-dimensional vector space in which semantically similar strings are located close to one another, allowing us to detect duplicate inventions that differ only in minor surface-level phrasing. Specifically, we use the all-MiniLM-L6-v2 sentence embedding model to represent each invention name as a vector of length $n = 384$.[28]

We then compute pairwise cosine similarity scores across all invention embedding vectors. The cosine similarity between two invention embedding vectors x and y is defined as:

$$\text{Cosine Similarity}(x, y) = \frac{\sum_{i=1}^{n} x_i \cdot y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2}}, \tag{A.1}$$

We then flag pairs of inventions as potential duplicates based on the cosine similarity as well as additional criteria from the metadata on the invention year, inventor, and the institution (e.g., Peenemünde Army Research Center, Intel, or University of Helsinki):

1. $Cosine\ Similarity(x, y) > 0.8$ and $Invention\ Year_x = Invention\ Year_y$
2. $Cosine\ Similarity(x, y) > 0.7$ and $Invention\ Year_x = Invention\ Year_y$ and $Inventor_x = Inventor_y$
3. $Cosine\ Similarity(x, y) > 0.7$ and $Invention\ Year_x = Invention\ Year_y$ and $Institution_x = Institution_y$

[28] The all-MiniLM-L6-v2" model is one of the most widely used sentence embedding models. For example, it was the third-most-downloaded model on huggingface.com as of July 2024. The findings do not depend on the choice of a specific model.

4. $Cosine\ Similarity(x, y) > 0.9$ and $Invention\ Year_x = Invention\ Year_y \pm 4$ and $Institution_x = Institution_y$

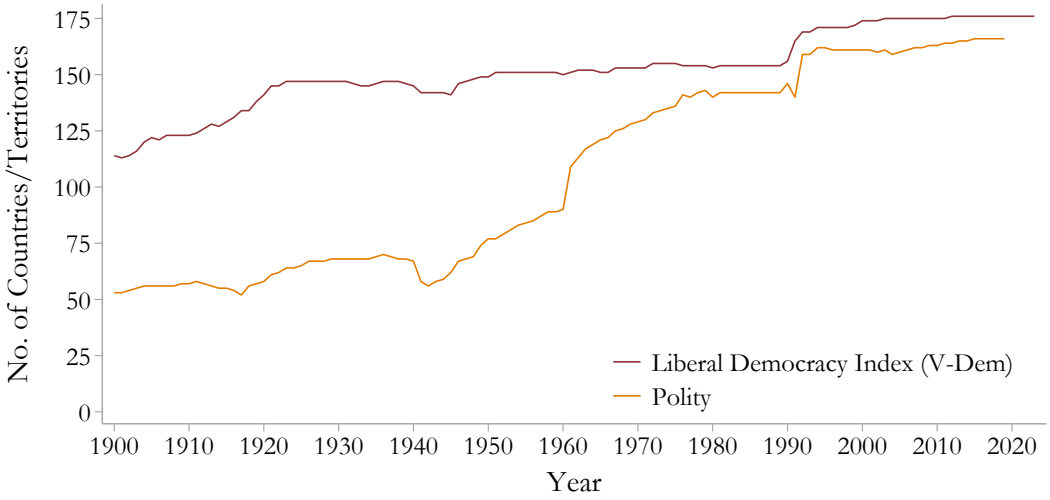
This procedure flags 4,903 inventions as duplicates. Manual inspection confirms that nearly all flagged pairs are substantively identical entries that correspond to only one distinct discovery. Table A.3 provides several illustrative examples of duplicate inventions identified by our procedure. We keep only one of these entries when we estimate our main results.

Table A.3: Examples of Duplicate Inventions

Group	Invention	Year	Inventor	Institution	Similarity
1	Cadmium-Nickel Battery	1899	Waldemar Jungner		0.995
	Nickel-Cadmium Battery	1899	Ernst Waldemar Jungner		
2	Polikarpov R-5	1928	Nikolai Polikarpov	Polikarpov Design Bureau	0.933
	Polikarpov I-5	1930	Nikolai Nikolaevich Polikarpov	TsKB-39	
3	Volksempfanger	1933	Otto Griessing	Reichs-Rundfunk-Gesellschaft	1.000
	Volksempfänger	1933	Otto Griessing	Reichspostzentralamt	
4	Ballistic Missile	1942	Wernher von Braun	Peenemünde Army Research Center	0.899
	Missile	1942	Wernher von Braun	Peenemünde Army Research Center	
5	Nuclear weapon	1945	Julius Robert Oppenheimer	Los Alamos Laboratory	0.700
	Pure fission weapon	1945	J. Robert Oppenheimer	Los Alamos Laboratory	
6	Computer processor	1971	Federico Faggin	Intel	0.700
	Intel microprocessor	1971	Federico Faggin	Intel	
7	Nintendo Color TV Game	1977	Masayuki Uemura	Nintendo	0.899
	Nintendo Color TV-Game consoles	1977	Masayuki Uemura	Nintendo Co., Ltd.	
8	Graphics workstations	1982	Jim Clark	Silicon Graphics	0.960
	Graphics workstation	1984	Jim Clark	Silicon Graphics, Inc.	
9	Linux (kernel)	1991	Linus Torvalds	University of Helsinki	0.985
	Kernel (Linux)	1991	Linus Torvalds	University of Helsinki	
10	MenACWY	2002		Sanofi Pasteur	0.700
	MenACWY vaccine	2002		Sanofi Pasteur	
11	Graphene	2004	Andre Geim	University of Manchester	0.802
	Graphene production techniques	2004	Andre Geim	University of Manchester	

A.3 Additional Details: Political Institutions

Figure A.2: Coverage of Political Institutions in V-Dem and Polity



Notes: The figure shows the number of countries/territories covered by the Varieties of Democracy project (V-Dem) and the Polity datasets for the period 1900–2020.

B Additional Results and Robustness on Section 3

B.1 Robustness

Table B.1: Political Institutions and the Size of the Academic Sector (Log/Asinh Transformation)

	(1)	(2)	(3)	(4)
Panel A: Dep. Var.: Log Universities (1900–2020)				
Institutional Quality	2.057***	1.719***	0.379*	0.431**
	(0.167)	(0.181)	(0.207)	(0.203)
Observations	968	968	968	968
R^2	0.702	0.742	0.889	0.892
Dependent Variable Mean	2.06	2.06	2.06	2.06
Panel B: Dep. Var.: Log Academics at Universities (1900–1969)				
Institutional Quality	3.527***	3.547***	0.865**	0.987**
	(0.364)	(0.369)	(0.408)	(0.392)
Observations	398	398	398	398
R^2	0.525	0.528	0.876	0.877
Dependent Variable Mean	5.33	5.33	5.33	5.33
Panel C: Dep. Var.: Asinh Universities (1900–2020)				
Institutional Quality	2.485***	1.870***	0.185	0.261
	(0.187)	(0.201)	(0.233)	(0.235)
Observations	1,343	1,343	1,343	1,343
R^2	0.719	0.763	0.891	0.894
Dependent Variable Mean	2.02	2.02	2.02	2.02
Panel D: Dep. Var.: Asinh Academics at Universities (1900–1969)				
Institutional Quality	7.143***	6.895***	1.012*	1.147**
	(0.529)	(0.517)	(0.570)	(0.558)
Observations	774	774	774	774
R^2	0.651	0.665	0.895	0.896
Dependent Variable Mean	3.10	3.10	3.10	3.10
Log / Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: Panel A reports estimates of Equation (1) with the log of the number of universities in country c and year t as the dependent variable. Panel B reports estimates of Equation (2) with the log of the number of academics as the dependent variable. Panel C reports estimates of Equation (1) with the inverse hyperbolic sine of the number of universities in country c and year t as the dependent variable. Panel D reports estimates of Equation (2) with the inverse hyperbolic sine of the number of academics as the dependent variable. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the log (panel a and b) or the inverse hyperbolic sine (panel c and d) of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

Table B.2: Political Institutions and the Size of the Academic Sector (Alternative Outcome: All Higher Education Institutions)

	Dep. Var.: Higher Education Institutions			
	(1)	(2)	(3)	(4)
Panel A: All Institutions (1900–2020)				
Institutional Quality	150.9***	130.1**	112.1***	105.6***
	(52.07)	(58.28)	(42.70)	(40.34)
Observations	1,343	1,343	1,343	1,343
R^2	0.302	0.306	0.671	0.672
Dependent Variable Mean	52.95	52.95	52.95	52.95
Panel B: Excluding "Other Higher Education Institutions" (1900–2020)				
Institutional Quality	141.1***	122.9**	103.3***	97.05***
	(50.02)	(56.40)	(39.39)	(37.13)
Observations	1,343	1,343	1,343	1,343
R^2	0.312	0.317	0.695	0.696
Dependent Variable Mean	49.71	49.71	49.71	49.71
Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of Equation (1). In Panel A, we report results that use the number of all higher education institutions (not just "main universities") in country c and year t as the dependent variable. In Panel B, we report results that use the number of all higher education institutions, excluding "other higher education institutions" in country c and year t as the dependent variable. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

Table B.3: Political Institutions and the Size of the Academic Sector - Alternative Backfilling Windows for University Counts

	Dep. Var.: Universities (1900–2020)											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Backfilling Window:	No Backfilling		20 Years		25 Years		30 Years		50 Years		120 Years	
Institutional Quality	74.63***	67.90***	76.81**	70.91***	74.87***	69.29***	74.90***	69.44***	71.68***	67.72***	62.92***	60.55***
	(28.27)	(24.74)	(29.52)	(26.51)	(28.42)	(25.75)	(28.21)	(25.62)	(25.23)	(23.71)	(22.45)	(21.85)
Observations	1,343	1,343	1,343	1,343	1,343	1,343	1,343	1,343	1,343	1,343	1,343	1,343
R^2	0.612	0.615	0.673	0.675	0.692	0.693	0.696	0.697	0.774	0.774	0.828	0.828
Dependent Variable Mean	25.69	25.69	28.71	28.71	28.97	28.97	29.18	29.18	31.88	31.88	34.21	34.21
Population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GDP per capita		Yes		Yes		Yes		Yes		Yes		Yes

Notes: The table reports estimates of variants of Equation (1) with the number of universities obtained by different backfilling windows in country c and year t as the dependent variable. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions include a control for population, time fixed effects, and country fixed effects. We additionally control for GDP per capita in even columns. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

C Additional Results and Robustness on Section 4

Table C.1: Political Institutions and Publications (Log Transformation)

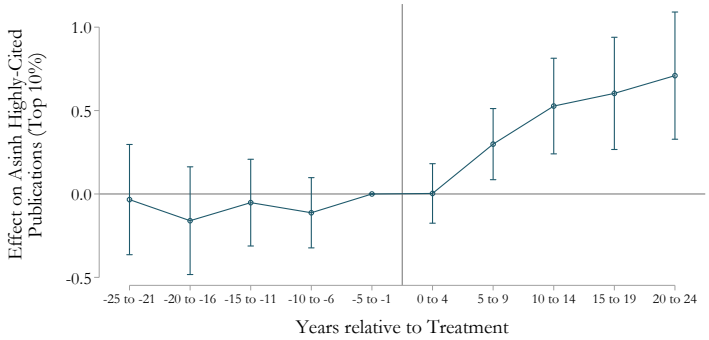
	Dep. Var.: Log Publications (1900–2020)			
	(1)	(2)	(3)	(4)
Institutional Quality	7.226***	5.682***	1.029***	0.978***
	(0.265)	(0.322)	(0.294)	(0.288)
Log Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes
Observations	3,288	3,288	3,288	3,288
R^2	0.678	0.794	0.944	0.944
Dependent Variable Mean	5.37	5.37	5.37	5.37

Notes: The table reports estimates of Equation (3), where the dependent variable is the log number of publications in country c within a ± 2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the log of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

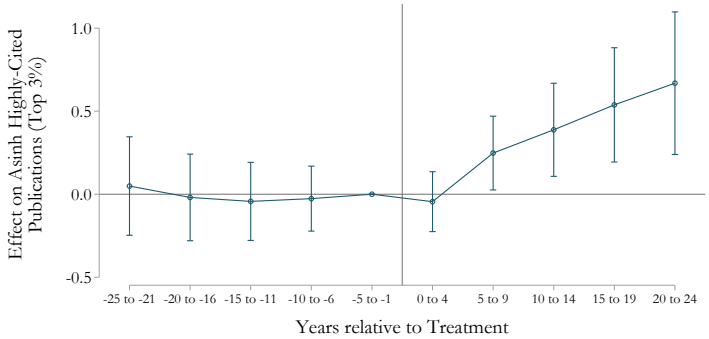
C.1 Additional Results on Sharp Changes in Political Institutions

Figure C.1: Event Study Estimates for Number of Highly-Cited Publications - Alternative Measures of Highly-Cited Publications

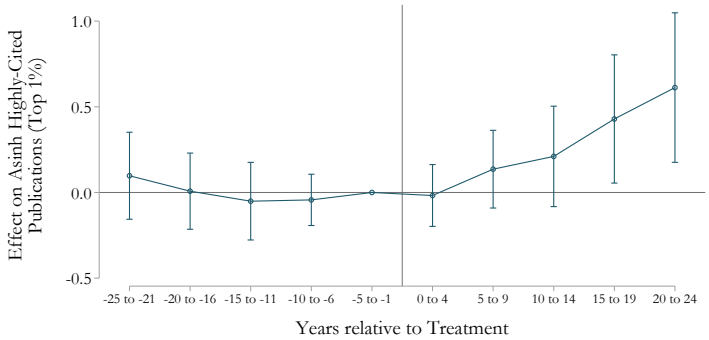
(a) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 10%)



(b) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 3%)



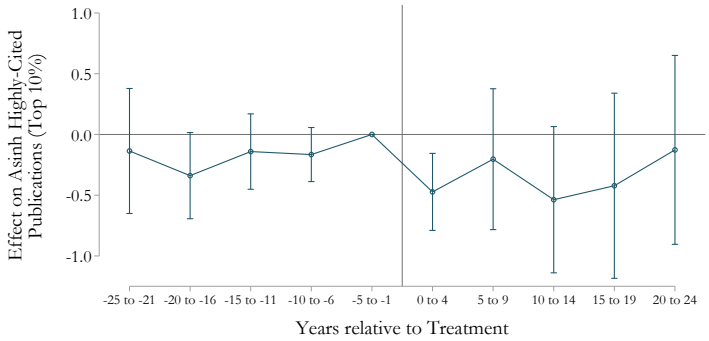
(c) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 1%)



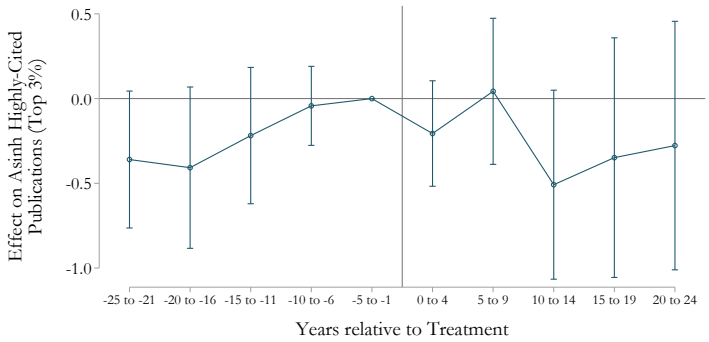
Notes: The figure plots the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5). The dependent variable is the inverse hyperbolic sine of the number of highly-cited publications in country c within a ± 2-year window around year t. Highly-cited publications are defined as publications in the top 10% (panel a), top 3% (panel b), or top 1% (panel c) of the discipline and year-level citation distribution.

Figure C.2: Event Study Estimates for Number of Highly-Cited Publications - Alternative Measures of Highly-Cited Publications

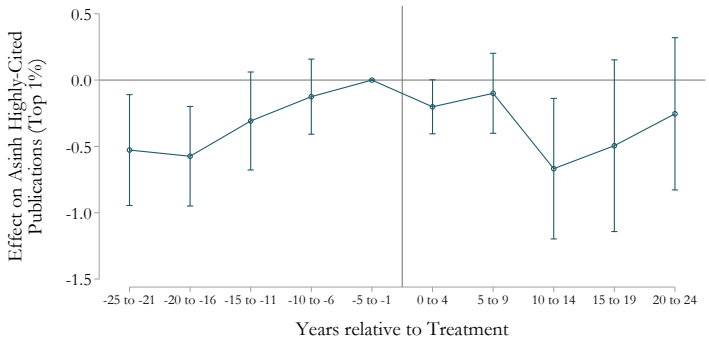
(a) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 10%)



(b) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 3%)



(c) Classic TWFE Event Study: Asinh Highly-Cited Publications (Top 1%)

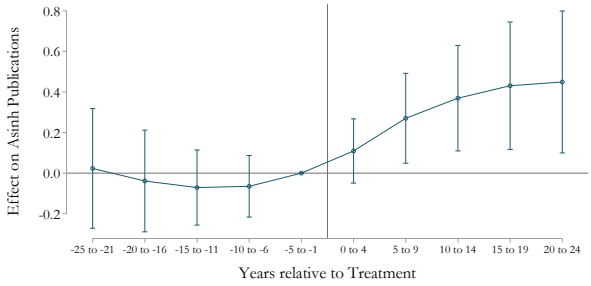


Notes: The figure plots the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (6). The dependent variable is the inverse hyperbolic sine of the number of highly-cited publications in country c within a ± 2-year window around year t. Highly-cited publications are defined as publications in the top 10% (panel a), top 3% (panel b), or top 1% (panel c) of the discipline and year-level citation distribution.

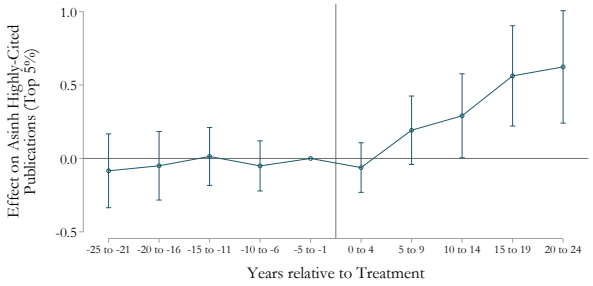
Figure C.3: Event Study Estimates – Alternative Threshold for Major Institutional Shock: 0.15

Positive Shocks

(a) Asinh Publications

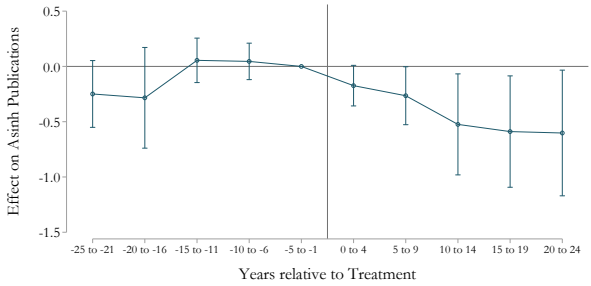


(b) Asinh Highly-Cited Publications (Top 5%)

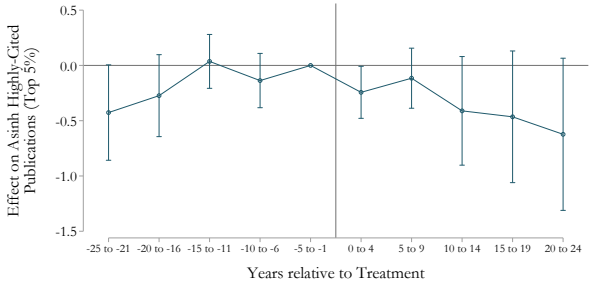


Negative Shocks

(c) Asinh Publications

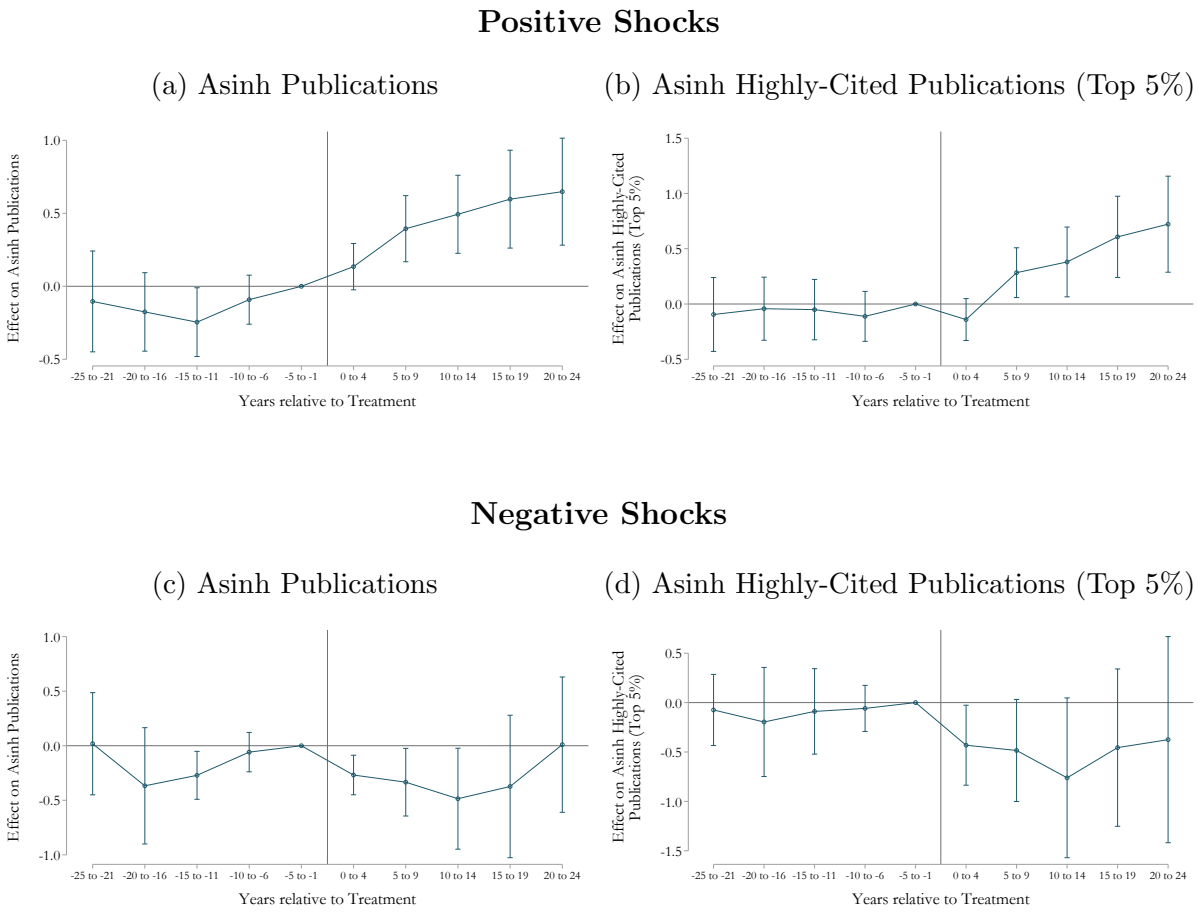


(d) Asinh Highly-Cited Publications (Top 5%)



Notes: Panels (a) and (b) plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5). The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited (top 5%) publications (panel b) in country c within a $\pm$2-year window around year t. Panels (c) and (d) plot equivalent estimates of Equation (6). Major institutional shocks are defined as episodes with LDI changes of at least 0.15 points within a five-year window.

Figure C.4: Event Study Estimates – Alternative Threshold for Major Institutional Shock: 0.25

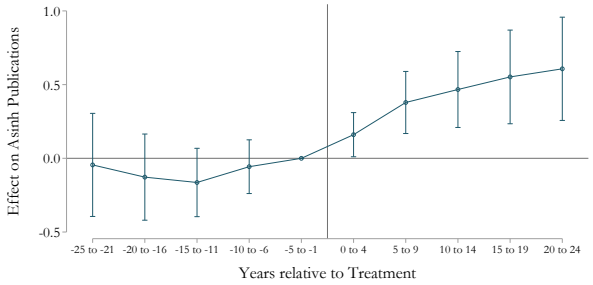


Notes: Panels (a) and (b) plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5). The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited (top 5%) publications (panel b) in country c within a $\pm$2-year window around year t. Panels (c) and (d) plot equivalent estimates of Equation (6). Major institutional shocks are defined as episodes with LDI changes of at least 0.25 points within a five-year window.

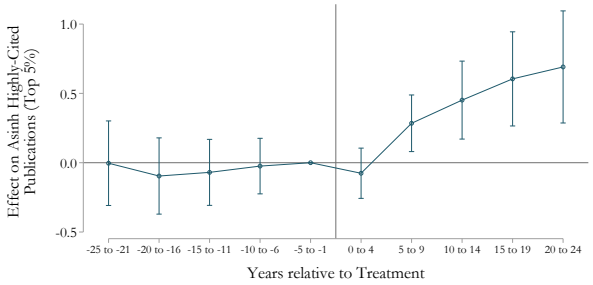
Figure C.5: Event Study Estimates – Controlling for Interstate Wars

Positive Shocks

(a) Asinh Publications

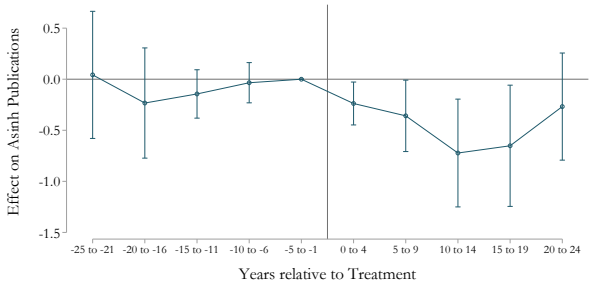


(b) Asinh Highly-Cited Publications (Top 5%)

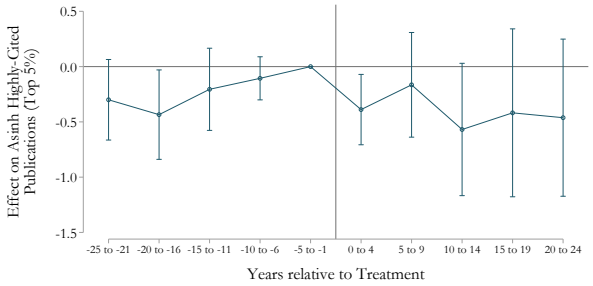


Negative Shocks

(c) Asinh Publications



(d) Asinh Highly-Cited Publications (Top 5%)

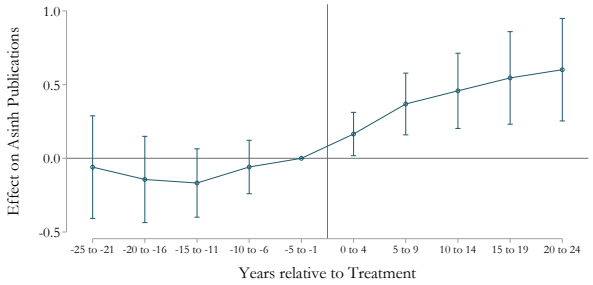


Notes: Panels (a) and (b) plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5). The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited (top 5%) publications (panel b) in country c within a $\pm$2-year window around year t. Panels (c) and (d) plot equivalent estimates of Equation (6). For all estimates, we control for battle deaths in interstate wars using data on interstate wars from the Correlates of War Project (Sarkees and Wayman, 2010) (see Section 4.2 for details).

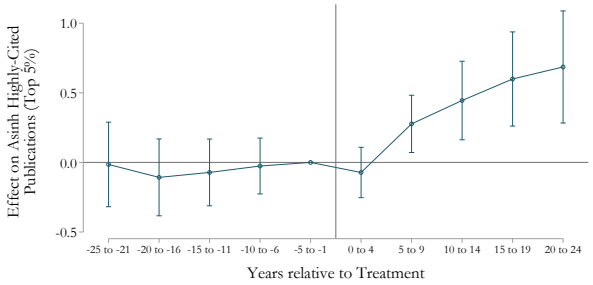
Figure C.6: Event Study Estimates – Controlling for Intrastate Wars

Positive Shocks

(a) Asinh Publications

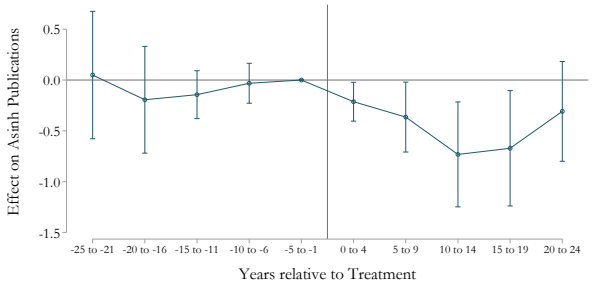


(b) Asinh Highly-Cited Publications (Top 5%)

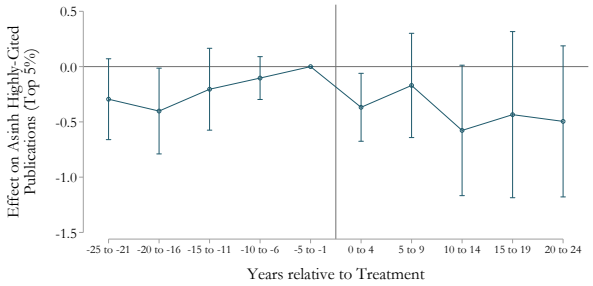


Negative Shocks

(c) Asinh Publications



(d) Asinh Highly-Cited Publications (Top 5%)

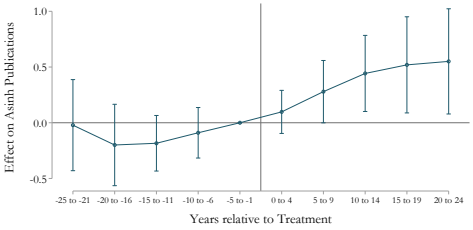


Notes: Panels (a) and (b) plot the point estimates and 95% confidence intervals of the TWFE event study coefficients outlined in Equation (5). The dependent variable is the inverse hyperbolic sine of the number of publications (panel a) or the number of highly-cited (top 5%) publications (panel b) in country c within a $\pm$2-year window around year t. Panels (c) and (d) plot equivalent estimates of Equation (6). For all estimates, we control for battle deaths in intrastate wars using data on intrastate wars from the Correlates of War Project (Sarkees and Wayman, 2010) (see Section 4.2 for details).

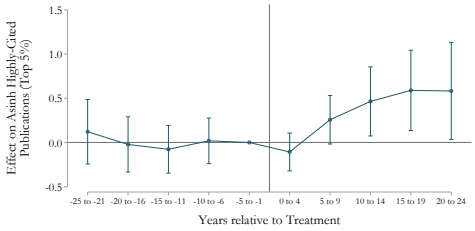
Figure C.7: Event Studies Positive Shocks – Omitting One Continent

Without Europe

(a) Asinh Publications

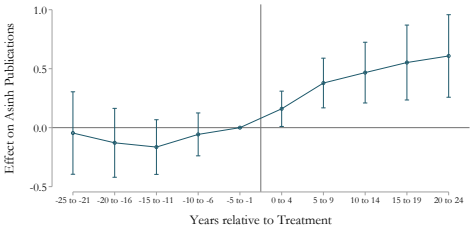


(b) Asinh Highly-Cited Pubs (Top 5%)

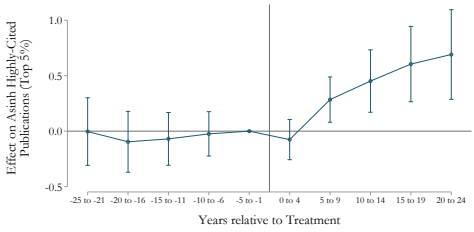


Without America

(c) Asinh Publications

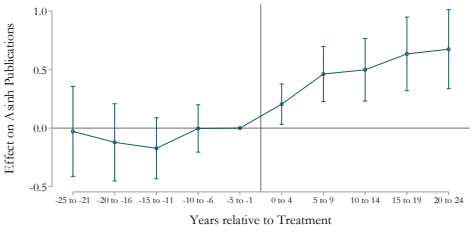


(d) Asinh Highly-Cited Pubs (Top 5%)

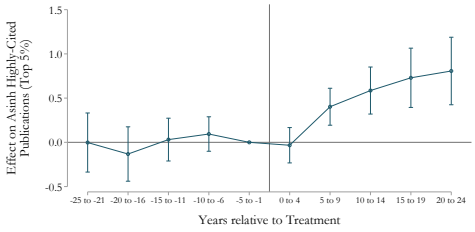


Without Asia

(e) Asinh Publications

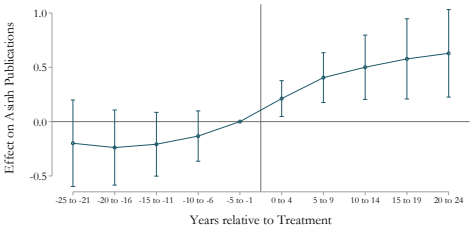


(f) Asinh Highly-Cited Pubs (Top 5%)

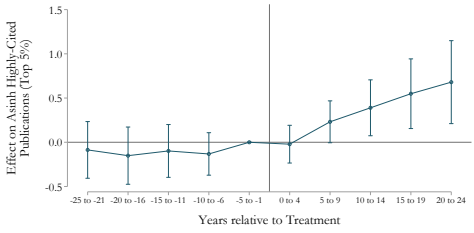


Without Africa

(g) Asinh Publications

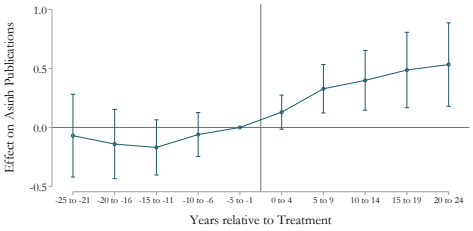


(h) Asinh Highly-Cited Pubs (Top 5%)

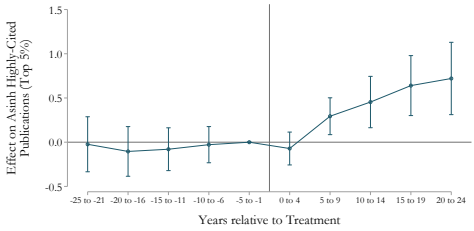


Without Oceania

(i) Asinh Publications



(j) Asinh Highly-Cited Pubs (Top 5%)

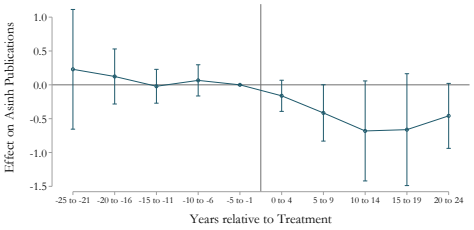


Notes: The figures plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (5). The dependent variable is the inverse hyperbolic sine of the number of publications (left column) or the number of highly-cited (top 5%) publications (right column) in country c within a $\pm$2-year window around year t. In each row, we omit one continent at a time. For example, in the first row, we omit all European countries (see Section 4.2 for details).

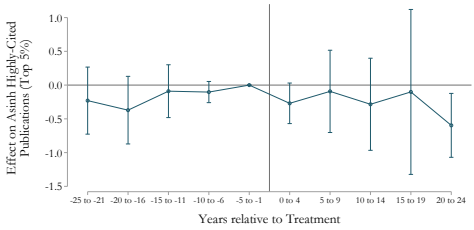
Figure C.8: Event Studies Negative Shocks – Omitting One Continent

Without Europe

(a) Asinh Publications

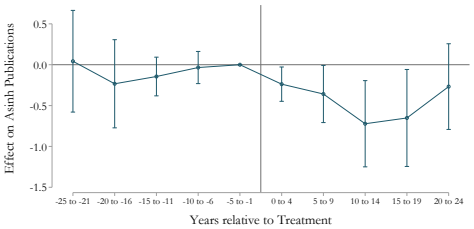


(b) Asinh Highly-Cited Pubs (Top 5%)

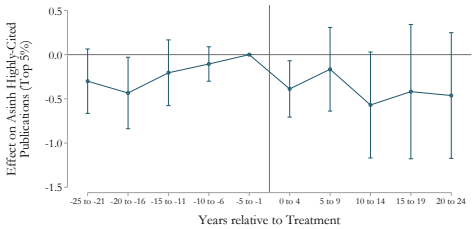


Without America

(c) Asinh Publications

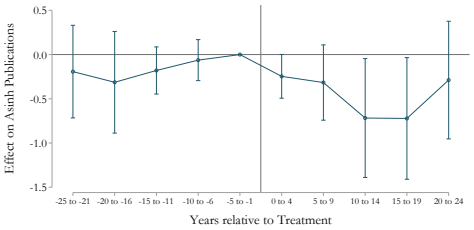


(d) Asinh Highly-Cited Pubs (Top 5%)

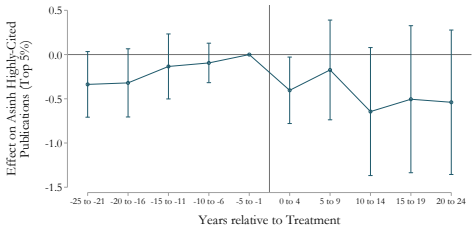


Without Asia

(e) Asinh Publications

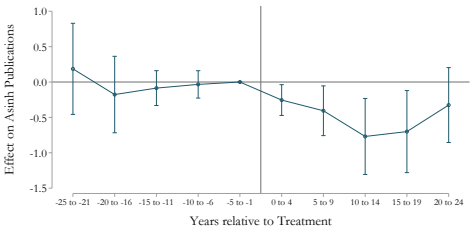


(f) Asinh Highly-Cited Pubs (Top 5%)

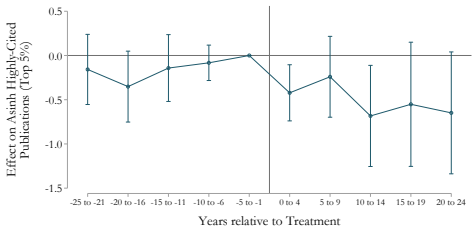


Without Africa

(g) Asinh Publications

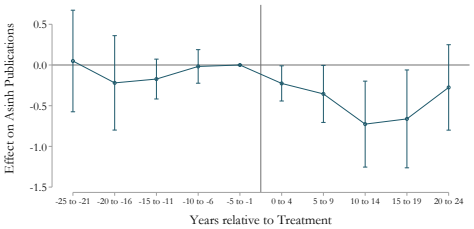


(h) Asinh Highly-Cited Pubs (Top 5%)

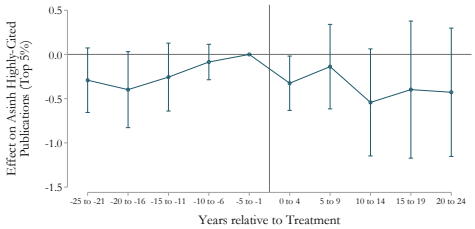


Without Oceania

(i) Asinh Publications



(j) Asinh Highly-Cited Pubs (Top 5%)



Notes: The figures plot the point estimates and 95% confidence intervals of the TWFE event study coefficients from Equation (6). The dependent variable is the inverse hyperbolic sine of the number of publications (left column) or the number of highly-cited (top 5%) publications (right column) in country c within a $\pm$2-year window around year t. In each row, we omit one continent. For example, in the first row, we omit all European countries (see Section 4.2 for details).

D Additional Results and Robustness on Section 6

D.1 Robustness: Alternative Cardinality Values

Table D.1: Political Institutions and Concentration of Ideas (Cardinality of 20)

	Dep. Var.: Concentration Ratio^{20}			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.408***	-0.314***	-0.008	-0.019
	(0.023)	(0.025)	(0.032)	(0.031)
Observations	3,288	3,288	3,288	3,288
R^2	0.533	0.646	0.828	0.834
Dependent Variable Mean	0.67	0.67	0.67	0.67
Panel B: Top 10% Research				
Institutional Quality	-0.449***	-0.383***	-0.040	-0.042
	(0.021)	(0.027)	(0.035)	(0.035)
Observations	2,625	2,625	2,625	2,625
R^2	0.491	0.589	0.818	0.818
Dependent Variable Mean	0.76	0.76	0.76	0.76
Panel C: Top 5% Research				
Institutional Quality	-0.452***	-0.399***	-0.062*	-0.061*
	(0.022)	(0.027)	(0.036)	(0.036)
Observations	2,401	2,401	2,401	2,401
R^2	0.481	0.578	0.814	0.814
Dependent Variable Mean	0.79	0.79	0.79	0.79
Panel D: Top 3% Research				
Institutional Quality	-0.459***	-0.412***	-0.093**	-0.092**
	(0.023)	(0.028)	(0.038)	(0.038)
Observations	2,252	2,252	2,252	2,252
R^2	0.479	0.572	0.814	0.815
Dependent Variable Mean	0.81	0.81	0.81	0.81
Panel E: Top 1% Research				
Institutional Quality	-0.449***	-0.419***	-0.149***	-0.148***
	(0.023)	(0.027)	(0.039)	(0.039)
Observations	1,916	1,916	1,916	1,916
R^2	0.479	0.553	0.793	0.803
Dependent Variable Mean	0.85	0.85	0.85	0.85
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 20. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.2: Political Institutions and Concentration of Ideas (Cardinality of 50)

	Dep. Var.: Concentration Ratio^{50}			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.264***	-0.206***	-0.022	-0.024
	(0.013)	(0.015)	(0.019)	(0.019)
Observations	3,288	3,288	3,288	3,288
R^2	0.556	0.668	0.835	0.836
Dependent Variable Mean	0.86	0.86	0.86	0.86
Panel B: Top 10% Research				
Institutional Quality	-0.254***	-0.221***	-0.048**	-0.045**
	(0.013)	(0.015)	(0.020)	(0.020)
Observations	2,625	2,625	2,625	2,625
R^2	0.498	0.592	0.809	0.815
Dependent Variable Mean	0.91	0.91	0.91	0.91
Panel C: Top 5% Research				
Institutional Quality	-0.248***	-0.224***	-0.064***	-0.061***
	(0.013)	(0.015)	(0.020)	(0.020)
Observations	2,401	2,401	2,401	2,401
R^2	0.493	0.573	0.795	0.807
Dependent Variable Mean	0.92	0.92	0.92	0.92
Panel D: Top 3% Research				
Institutional Quality	-0.241***	-0.221***	-0.071***	-0.069***
	(0.013)	(0.015)	(0.020)	(0.020)
Observations	2,252	2,252	2,252	2,252
R^2	0.487	0.556	0.786	0.803
Dependent Variable Mean	0.93	0.93	0.93	0.93
Panel E: Top 1% Research				
Institutional Quality	-0.210***	-0.198***	-0.062***	-0.061***
	(0.014)	(0.015)	(0.019)	(0.017)
Observations	1,916	1,916	1,916	1,916
R^2	0.454	0.511	0.761	0.790
Dependent Variable Mean	0.95	0.95	0.95	0.95
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 50. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

D.2 Robustness: Minimum Number of Publications

Table D.3: Institutions and Concentration of Ideas (Minimum 250 Publications)

	Dep. Var.: Concentration Ratio[25]			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.111***	-0.110***	-0.019	-0.017
	(0.013)	(0.014)	(0.019)	(0.019)
Observations	1,538	1,538	1,538	1,538
R^2	0.182	0.221	0.723	0.728
Dependent Variable Mean	0.55	0.55	0.55	0.55
Panel B: Top 10% Research				
Institutional Quality	-0.281***	-0.272***	-0.044	-0.040
	(0.021)	(0.024)	(0.040)	(0.040)
Observations	1,536	1,536	1,536	1,536
R^2	0.305	0.385	0.755	0.759
Dependent Variable Mean	0.67	0.67	0.67	0.67
Panel C: Top 5% Research				
Institutional Quality	-0.357***	-0.345***	-0.075*	-0.072*
	(0.023)	(0.027)	(0.043)	(0.043)
Observations	1,535	1,535	1,535	1,535
R^2	0.354	0.455	0.787	0.789
Dependent Variable Mean	0.73	0.73	0.73	0.73
Panel D: Top 3% Research				
Institutional Quality	-0.399***	-0.386***	-0.101**	-0.100**
	(0.025)	(0.028)	(0.042)	(0.042)
Observations	1,531	1,531	1,531	1,531
R^2	0.391	0.494	0.816	0.816
Dependent Variable Mean	0.77	0.77	0.77	0.77
Panel E: Top 1% Research				
Institutional Quality	-0.419***	-0.409***	-0.118***	-0.123***
	(0.024)	(0.028)	(0.033)	(0.033)
Observations	1,492	1,492	1,492	1,492
R^2	0.449	0.537	0.830	0.834
Dependent Variable Mean	0.84	0.84	0.84	0.84
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 25. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. We drop all country-year observations with fewer than 250 publications. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.4: Institutions and Concentration of Ideas (Minimum 500 Publications)

	Dep. Var.: Concentration Ratio[25]			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.092***	-0.092***	-0.008	-0.005
	(0.013)	(0.014)	(0.017)	(0.018)
Observations	1,287	1,287	1,287	1,287
R^2	0.147	0.183	0.712	0.716
Dependent Variable Mean	0.53	0.53	0.53	0.53
Panel B: Top 10% Research				
Institutional Quality	-0.225***	-0.223***	-0.046	-0.041
	(0.020)	(0.022)	(0.032)	(0.032)
Observations	1,287	1,287	1,287	1,287
R^2	0.276	0.333	0.762	0.767
Dependent Variable Mean	0.63	0.63	0.63	0.63
Panel C: Top 5% Research				
Institutional Quality	-0.308***	-0.305***	-0.105**	-0.099**
	(0.022)	(0.025)	(0.041)	(0.041)
Observations	1,286	1,286	1,286	1,286
R^2	0.328	0.406	0.765	0.769
Dependent Variable Mean	0.68	0.68	0.68	0.68
Panel D: Top 3% Research				
Institutional Quality	-0.368***	-0.365***	-0.135***	-0.130***
	(0.024)	(0.027)	(0.043)	(0.044)
Observations	1,286	1,286	1,286	1,286
R^2	0.364	0.453	0.792	0.794
Dependent Variable Mean	0.73	0.73	0.73	0.73
Panel E: Top 1% Research				
Institutional Quality	-0.418***	-0.416***	-0.140***	-0.144***
	(0.025)	(0.029)	(0.035)	(0.035)
Observations	1,277	1,277	1,277	1,277
R^2	0.430	0.518	0.832	0.833
Dependent Variable Mean	0.81	0.81	0.81	0.81
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 25. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. We drop all country-year observations with fewer than 500 publications. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.5: Institutions and Concentration of Ideas (Minimum 1,000 Publications)

	Dep. Var.: Concentration Ratio[25]			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.081***	-0.083***	-0.011	-0.008
	(0.014)	(0.014)	(0.017)	(0.017)
Observations	1,052	1,052	1,052	1,052
R^2	0.142	0.163	0.738	0.739
Dependent Variable Mean	0.52	0.52	0.52	0.52
Panel B: Top 10% Research				
Institutional Quality	-0.179***	-0.183***	-0.040	-0.032
	(0.019)	(0.020)	(0.025)	(0.026)
Observations	1,052	1,052	1,052	1,052
R^2	0.253	0.285	0.768	0.772
Dependent Variable Mean	0.60	0.60	0.60	0.60
Panel C: Top 5% Research				
Institutional Quality	-0.248***	-0.255***	-0.098***	-0.087**
	(0.022)	(0.023)	(0.036)	(0.037)
Observations	1,052	1,052	1,052	1,052
R^2	0.303	0.355	0.763	0.769
Dependent Variable Mean	0.64	0.64	0.64	0.64
Panel D: Top 3% Research				
Institutional Quality	-0.308***	-0.318***	-0.138***	-0.124***
	(0.024)	(0.026)	(0.045)	(0.046)
Observations	1,052	1,052	1,052	1,052
R^2	0.329	0.400	0.777	0.782
Dependent Variable Mean	0.68	0.68	0.68	0.68
Panel E: Top 1% Research				
Institutional Quality	-0.395***	-0.408***	-0.147***	-0.145***
	(0.026)	(0.029)	(0.041)	(0.043)
Observations	1,050	1,050	1,050	1,050
R^2	0.397	0.481	0.826	0.826
Dependent Variable Mean	0.77	0.77	0.77	0.77
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: *Notes:* The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 25. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. We drop all country-year observations with fewer than 1,000 publications. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.6: Institutions and Concentration of Ideas (Minimum 2,000 Publications)

	Dep. Var.: Concentration Ratio[25]			
	(1)	(2)	(3)	(4)
Panel A: All Research				
Institutional Quality	-0.077***	-0.080***	-0.026*	-0.023
	(0.014)	(0.015)	(0.016)	(0.016)
Observations	850	850	850	850
R^2	0.151	0.186	0.789	0.791
Dependent Variable Mean	0.51	0.51	0.51	0.51
Panel B: Top 10% Research				
Institutional Quality	-0.154***	-0.160***	-0.039*	-0.034
	(0.019)	(0.019)	(0.020)	(0.021)
Observations	850	850	850	850
R^2	0.257	0.288	0.797	0.800
Dependent Variable Mean	0.57	0.57	0.57	0.57
Panel C: Top 5% Research				
Institutional Quality	-0.205***	-0.213***	-0.083***	-0.074**
	(0.021)	(0.021)	(0.029)	(0.030)
Observations	850	850	850	850
R^2	0.315	0.354	0.785	0.789
Dependent Variable Mean	0.60	0.60	0.60	0.60
Panel D: Top 3% Research				
Institutional Quality	-0.259***	-0.269***	-0.125***	-0.114***
	(0.023)	(0.024)	(0.037)	(0.038)
Observations	850	850	850	850
R^2	0.342	0.391	0.781	0.786
Dependent Variable Mean	0.63	0.63	0.63	0.63
Panel E: Top 1% Research				
Institutional Quality	-0.368***	-0.384***	-0.169***	-0.164***
	(0.027)	(0.028)	(0.045)	(0.046)
Observations	849	849	849	849
R^2	0.383	0.463	0.800	0.801
Dependent Variable Mean	0.72	0.72	0.72	0.72
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes

Notes: The table reports estimates of Equation (7), where the dependent variable is the concentration ratio of publications or highly-cited publications in country c within a $\pm$2-year window around year t with a cardinality k of 25. Panel A shows results for all publications. Panel B shows results for publications in the top 10%, Panel C for the top 5%, Panel D for the top 3%, and Panel E for the top 1% of the field and year-level citation distribution. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. We drop all country-year observations with fewer than 2,000 publications. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

E Additional Results and Robustness on Section 7

Table E.1: Political Institutions and Technological Breakthroughs – without Deduplication

	Dep. Var.: Asinh Discoveries			
	(1)	(2)	(3)	(4)
Institutional Quality	1.884***	2.486***	0.377**	0.364**
	(0.310)	(0.364)	(0.162)	(0.158)
Asinh Population	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes
Country FE			Yes	Yes
GDP per capita				Yes
Observations	3,774	3,774	3,774	3,774
R^2	0.315	0.414	0.863	0.863
Dependent Variable Mean	0.45	0.45	0.45	0.45

Notes: The table reports estimates of Equation (8), where the dependent variable is the inverse hyperbolic sine of the number of technological breakthroughs in country c within a ± 2-year window around year t. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for the inverse hyperbolic sine of population. In column (2), we additionally control for time fixed effects; in column (3), we add country fixed effects; and in column (4), we further control for GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.2: Political Institutions and Technological Breakthroughs (Counts)

	Dep. Var.: Discoveries (1900–2020)			
	All	Winsorized 97.5pct	Winsorized 95pct	Winsorized 90pct
	(1)	(2)	(3)	(4)
Institutional Quality	12.036	1.940*	0.798**	0.242**
	(10.293)	(1.102)	(0.325)	(0.120)
Population	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
GDP per capita	Yes	Yes	Yes	Yes
Observations	3,774	3,774	3,774	3,774
R^2	0.715	0.841	0.831	0.789
Dependent Variable Mean	5.54	1.89	1.02	0.41

Notes: The table reports estimates of Equation (8), where the dependent variable is the number of technological breakthroughs. In columns (2)-(4), we present estimates that use the winsorized number of discoveries as the dependent variable. The column headings indicate the winsorization percentile for each specification. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for population, time fixed effects, country fixed effects, and GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.3: Technological Breakthroughs by Institution Type (Counts)

	Dep. Var.: Discoveries (1900–2020) by Institution Type				
	Universities	Government	Military	Industry	Independent
	(1)	(2)	(3)	(4)	(5)
Institutional Quality	0.686**	0.317*	-0.078	0.870	0.024
	(0.314)	(0.191)	(0.052)	(0.552)	(0.056)
Population	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
GDP per capita	Yes	Yes	Yes	Yes	Yes
Observations	3,774	3,774	3,774	3,774	3,774
R^2	0.735	0.723	0.532	0.812	0.545
Dependent Variable Mean	0.44	0.31	0.03	0.74	0.09

Notes: The table reports estimates of Equation (8), where the dependent variable is the number of technological breakthroughs by institution type in country c within a ±2-year window around year t which we winsorize at the 97.5 percentile. The main explanatory variable is country c's V-Dem Liberal Democracy Index in year $t-1$. All regressions control for population, time fixed effects, country fixed effects, and GDP per capita. Standard errors are clustered at the country level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.