The Spatial Concentration of Innovation over the (Very) Long Run*

Francisco Buera† Martin Kornejew‡ Yueran Ma§
Ezra Oberfield¶ Nicholas Trachter‖

March 2026

Abstract

We analyze significant technological innovations over the past one thousand years by constructing a dataset from Wikipedia titles, which records the time, location, inventor type, domain of application, among other background information about the innovations. The dataset reveals several striking facts. Across space, innovations are highly concentrated at the country or continent level, much more than population or output, even among high-income countries. Over time, the center of innovation shifts geographically over the span of centuries. The level of innovations per capita has decreased relative to one century ago in many European countries, and increased steadily in the United States since the founding of the country. We then use the data to inform theories of technological change with spatial agglomeration and diffusion. The vast heterogeneity in innovations per capita across space and over time even among countries with similar incomes highlights that innovations are not just a function of population, and points to the existence of some elastic factor that accelerates innovation when local conditions are favorable. This factor could be one that is mobile across countries, such as financial or human capital, or mobile within a country between innovation and adoption. Technology diffusion then allows for advanced countries to have similar incomes even when innovation is highly concentrated. Some suggestive evidence points to financial capital, which shows a positive correlation with a country's innovation intensity.

*We thank Jingyi Jia and Howard Guankun Li for extraordinary research assistance.
†Washington University in St. Louis and NBER (fjbuera@wustl.edu).
‡Bocconi University (martin.kornejew@unibocconi.it).
§University of Chicago Booth School of Business and NBER (yueran.ma@chicagobooth.edu).
¶Cornell University and NBER (ezraoberfield@cornell.edu).
‖Federal Reserve Bank of Richmond (trachter@gmail.com).

1 Introduction

Technological innovations are fundamental for economic growth and development. Yet we do not have a comprehensive understanding about their patterns of emergence, across different parts of the world and over a long period of time. Where do significant innovations come from during the past millennium, and how concentrated are they? Is the volume of innovation proportional to population, or are some places disproportionately innovative? Do centers of innovation persist, or do they rise and fall? And how much do innovation activities and their concentration translate into concentration of output?

These questions have been difficult to answer on a global scale over the long run, in large part due to data constraints. Patent data, the most commonly used measure of inventive activity, can provide information for specific countries in recent decades, but are difficult to compare across countries and over time. Similarly, R&D expenditures and researcher counts are available only for recent decades and a subset of countries. Economic historians have produced rich case studies of particular episodes—the Industrial Revolution in Britain, the Golden Age of Islamic science, the rise of American invention—but broad-based, long-run quantitative analyses of innovations have been limited.

In this paper, we assemble a dataset of significant technological innovations over the past one thousand years, and use it to document a set of facts about the emergence of innovations around the world in the long run. We then develop a simple framework with spatial agglomeration and technology diffusion to interpret these facts, with particular attention to the role of mobile factors in shaping the concentration of innovative activity.

Our dataset is built from the universe of English Wikipedia titles. Following the insights of Asirvatham (2024) and Kalyani et al. (2025), we use large language models to filter Wikipedia titles for those describing technological innovations, and extract their background information such as the time and place of invention, the type of inventor, and the domain of application. After filtering and validation, the final dataset records 10,375 innovations spanning the years 1000 to 2000 AD. The data capture the canonical innovations of successive eras (e.g.,the Newcomen atmospheric engine, the spinning jenny, the Watt steam engine, the internal combustion engine, the incandescent light bulb, the airplane, the transistor, the microprocessor, and the mobile phone). We cross-check the Wikipedia-based measure against patent data in recent decades, and verify that the two types of data align with each other.

The data reveal several striking facts. First, innovation is highly concentrated geographically, much more than population or economic output. The Herfindahl-Hirschman Index for innovation lies well

above the corresponding index for population throughout the sample, at the continent, country, and U.S. state levels. This gap persists even among high-income countries. This concentration of innovation even within rich countries stands in contrast to a simple narrative in which countries at the frontier innovate while the countries behind the frontier imitate. Interestingly, artwork such as literature (a genre that has a long history and remains popular over time) is also spatially concentrated (which we measure using a parallel Wikipedia dataset), but less so than technological innovations. The extreme concentration appears to be a distinctive feature of technology, not a generic property of intellectual output.

Second, the geographic center of innovation shifts over the span of centuries. In the early part of our sample, innovative activity was more dispersed, with notable contributions from Asia and the Middle East. Europe emerged as the dominant center during the Scientific Revolution and maintained this position through the Industrial Revolution. The United States then rose steadily and became the single largest source of significant innovations by the early twentieth century. These are not gradual adjustments—they are dramatic reallocations of the world's inventive capacity.

Third, innovations per capita exhibits striking patterns of rise and decline. For example, the United Kingdom had very high innovations per capita during and after the Industrial Revolution, but its innovation intensity has fallen markedly since then—even though the country did not experience obvious deterioration in institutional quality. A similar pattern holds for several other European countries. Meanwhile, the United States has experienced a steady increase in innovations per capita since the founding of the country, a trajectory sustained for more than two centuries. The heterogeneity in innovations per capita across countries—even among those with similar incomes—is one noteworthy feature of the data.

Fourth, the spatial correlation of innovation changes sign over time. Before 1900, innovation growth in one country was positively associated with innovation growth in nearby countries, consistent with knowledge spillovers. After 1900, this correlation turns negative: when a country innovates faster, its neighbors tend to innovate more slowly. This sign reversal survives controls for trade linkages, common language, shared religion, colonial history, and bilateral trade intensity, and it appears within the United States across states as well. The pattern points to a mobile factor that flows toward innovating locations, generating gains in one place partly at the expense of nearby ones.

Fifth, the concentration of innovation is broad-based across technological fields rather than driven by the dominance of any particular domain. A decomposition into global domain-specific and local country-specific components shows that the local component accounts for a substantially larger share

of the variance in innovation growth. Countries do not rise and fall as innovators because they happen to specialize in globally ascendant technologies. They rise and fall because of something local.

These facts, taken together, are informative about the forces shaping the geography of innovation. The heterogeneity in innovations per capita—even among countries with similar populations and incomes—suggests that innovations are not simply a function of scale. The decline of innovation in countries with stable institutions sits uncomfortably with purely institutional accounts. And the shift from positive to negative spatial correlation suggests that in the modern era, the flow of a mobile factor has become an increasingly important force relative to local agglomeration spillovers.

What is the mobile factor? Our framework features locations that differ in their capacity for research, adoption, and production. Research generates new techniques; technology diffusion allows countries to benefit from innovations produced elsewhere, which helps explain why countries with similar incomes can have quite different levels of innovation activity. But the concentration of innovation requires some elastic input that accelerates inventive activity when local conditions are favorable. This factor could be financial capital, human capital, or resources reallocated within a country between innovation and adoption. Some suggestive evidence points to financial capital: across countries and over time, financial development shows a strong positive association with innovation intensity. The flow of financial resources toward locations with favorable conditions—and away from those without—can account for both the concentration of inventive activity and the negative spatial correlation in the modern era, when capital became more mobile.

Related literature This paper makes two main contributions. First, we provide new facts about the emergence of innovation over the very long run. A large body of work documents that innovation clusters geographically, driven by localized knowledge spillovers and agglomeration economies (Jaffe, Trajtenberg, and Henderson, 1993; Audretsch and Feldman, 1996; Akcigit et al., 2018; Moretti, 2021; Giroud, Liu, and Mueller, 2026; Chikis, Kleinman, and Prato, 2025); see Carlino and Kerr (2015) for a survey. This literature focuses primarily on within-country variation over recent decades. Over longer horizons, Bergeaud, Gozen, and Van Reenen (2026) measure cross-country "technological trajectories" from two centuries of patent data, documenting the relative decline of the UK and the rise of the U.S. and Germany. Akcigit, Grigsby, and Nicholas (2017) link historical U.S. patents to county-level data and find that innovation is associated with financial development and geographic connectedness. Comin, Easterly, and Gong (2010) push the horizon further back, showing that technology levels in 1500 AD predict income today. Our paper uses a different data source—Wikipedia titles processed with large language models (Asirvatham, Mokski, and Shleifer, 2026; Kalyani et al., 2025)—to trace innovation

concentration at the global level over a millennium. The collection of empirical facts we document suggest that innovation is not simply a function of scale or institutions, and point to the flow of a mobile factor as an increasingly important force shaping the geography of invention. An empirical literature documents how technologies trickle down from leaders to followers with substantial lags (Comin and Hobijn, 2004, 2010; Comin and Mestieri, 2018). Liu and Ma (2021) study how cross-sector knowledge spillovers through innovation networks shape the optimal allocation of R&D resources. A related body of work links finance, and equity markets in particular, to innovation (King and Levine, 1993; Brown, Fazzari, and Petersen, 2009; Hsu, Tian, and Xu, 2014; Comin and Nanda, 2019; see Kerr and Nanda, 2015 for a survey).

Second, we provide new theoretical insights about the forces behind innovation concentration. Our framework builds on models of idea-driven growth (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992) particularly the strand focusing on semiendogenous growth (Jones, 1995; Kortum, 1997). Our analysis also emphasizes international technology diffusion (Eaton and Kortum, 1999; Alvarez et al., 2013; Benhabib, Perla, and Tonetti, 2014; Buera and Oberfield, 2020; Cai, Li, and Santacreu, 2022; Lind and Ramondo (2023)), and in particular a tradeoff between innovation and adoption (Jovanovic and Rob, 1989; Acemoglu, Aghion, and Zilibotti, 2006; Trouvain, 2024). Diffusion can explain why countries with different innovation levels have similar incomes, but not why innovation is so concentrated in the first place. We introduce a mobile factor that flows toward favorable locations, generating concentration. Our analysis of the cross-country dynamics builds on Golub and Jackson, 2012, Cavalcanti, Giannitsarou, and Johnson, 2017, and Allen and Donaldson, 2020

2 Innovations Dataset from Wikipedia Titles

We aim to measure innovations around the world over a long period of time. To do so, we follow the insights of Asirvatham (2024) and Kalyani et al. (2025) to collect information from Wikipedia titles using large language models (LLMs).

2.1 Main Dataset for Technological Innovations

We obtain English Wikipedia titles from the Wikimedia English Wikipedia Dumps. First, we filter Wikipedia titles using LLMs to extract those about technological innovations, as explained in detail in Appendix IA1.1. Second, we use LLMs to compile information about these innovations, including

the time, location, inventor type (e.g., individuals, companies, governments, universities/nonprofits), domain of application (based on the structure of two-digit NAICS codes), as explained in detail in Appendix IA1.2.

After filtering the entire set of English Wikipedia titles for technological innovations and deduplication, we have 11,823 titles from 1000 AD to 2000 AD. Currently we use innovations after 1000 AD because innovations before then appear too sparse. We end in 2000 because some very recent innovations may not have been comprehensively recorded yet. We are not able to obtain location information for 1,448 titles. Since our analyses focus on spatial distribution of innovations, we have to drop those titles. The final dataset records 10,375 Wikipedia titles about technological innovations.

For each innovation, we collect background information from the Wikipedia page content using LLMs (see Appendix IA1.2). Specifically, we record the year when the innovation first became workable and the country where it was first created.[1] We also classify the type of inventor into four categories: individual inventors working independently of large organizations, private companies, government projects (including military initiatives), and universities or nonprofits. Moreover, we assign each innovation a domain of application from 15 categories following the structure of two-digit NAICS codes, supplemented with categories for military and defense and basic research. For innovations originating in the United States, we also collect additional information about the state of invention (see Appendix IA1.3 for details).

We check several key examples to verify that LLM filtering and information collection are reliable. For example, for the first industrial revolution, our dataset includes the Newcomen atmospheric engine (1712, United Kingdom), the spinning jenny (1764, United Kingdom), the Watt steam engine (1776, United Kingdom), the spinning mule (1779, United Kingdom), the power loom (1785, United Kingdom), and the steam locomotive (1804, United Kingdom). For the second industrial revolution, our dataset includes the Bessemer process (1856, United Kingdom), the internal combustion engine (1876, Germany), the incandescent light bulb (1879, United States), the automobile (1886, Germany), the X-ray (1895, Germany), the airplane (1903, United States), and the assembly line (1913, United States). For the third industrial revolution, our dataset includes the transistor (1947, United States), the integrated circuit (1958, United States), the personal computer (1970, United States), the microprocessor (1971, United States), and the mobile phone (1973, United States).

We also cross check the Wikipedia titles against patent data from PATSTAT, which contains biblio-

[1]We follow the Maddison Project database in defining and tracking countries through time in order to merge the database's estimates of population and real GDP per capita.

graphical data on patent applications filed at major patent offices around the world. We extract patent application counts from the Application table, which records each application's filing authority and filing year. For each country-year pair, we count the total number of patent applications, using the application authority code as a proxy for the country where the patent was filed. We then aggregate these country-level counts to the regional level using a mapping from ISO two-letter country codes to the 27 regions used in our dataset. This yields a region-year panel of patent application counts spanning from the late 18th century (the earliest records appear in 1782 for the British Isles) through 2020, covering approximately 84 million patent applications in total. Our comparison shows that the two series broadly co-move over time within regions, providing reassurance that our Wikipedia-based measure captures meaningful variation in innovative activity.

2.2 Additional Dataset for Artwork

To compare patterns of technological inventions with achievements in other domains, we also analyze Wikipedia titles on artwork. First, we filter Wikipedia titles using LLMs to extract those about artwork, as explained in detail in Appendix IA2.1. Second, we use LLMs to compile information about the artwork from the Wikipedia page content, including the time and location of composition, as explained in detail in Appendix IA2.2. Our empirical analyses in Section 3 focus on literature, since it is a genre with a long history and one that has remained consistently popular over time. Other genres can be less suitable for long-run analyses—for example, film has a short history while painting has become less prevalent due to technological change. As with technological innovations, for U.S. literature, we also collect additional information about the state most closely associated with the creation of each work (see Appendix IA2.3).

3 Empirical Facts

This section presents the main empirical findings on the spatial distribution of innovation over the past millennium. We organize the evidence around four sets of facts: (i) the spatial concentration of innovation relative to population and output, (ii) the evolution of innovations per capita across regions and countries, (iii) the spatial correlation of innovation activity, and (iv) heterogeneity across different types of innovation.

We present results at three levels of geographic aggregation. The continent level offers the broadest

view of global shifts in innovation activity and allows us to trace these shifts over the longest time horizon, from 1000 to 2000 AD. The country level reveals heterogeneity within continents—for instance, the distinct trajectories of the United Kingdom, Germany, and France within Europe—and covers the period from 1500 to 2000, when country-level attribution of innovations becomes more reliable. The U.S. state level, available from around 1800 onward, provides further evidence within a single country, which helps us examine whether the patterns observed across countries also manifest within national borders.

At each level, we measure innovation activity as the cumulative count of innovations over a trailing window: 100 years at the continent level, 50 years at the country level, and 20 years at the U.S. state level. The trailing window serves to smooth the lumpiness inherent in innovation counts, particularly in the earlier parts of the sample when innovations are sparse. The window length decreases with the level of geographic disaggregation: finer geographic units have shorter time coverage and correspondingly more observations per unit of time in the later periods, making shorter smoothing windows both feasible and desirable for capturing meaningful time variation.

3.1 Spatial Concentration of Innovation over Time

We begin by documenting the spatial concentration of innovation. A central finding is that innovation is highly concentrated geographically—much more than population or economic output—and that the geographic center of innovation has shifted substantially over the past millennium.

Continent level Figure 1 presents the distribution of innovations across continents over the period 1000–2000 AD. The top panel shows each continent's share of global innovations, measured as cumulative innovations over the preceding 100 years to smooth lumpiness in the early data. The figure reveals pronounced shifts in the geographic center of innovation over time. In the early centuries of our sample, innovation activity was more dispersed, with notable contributions from Asia and the Middle East. Starting around the period of the Scientific Revolution, Europe emerged as the dominant center of innovation, a position it maintained through the First Industrial Revolution. Subsequently, North America—driven primarily by the United States—rose to become the leading source of significant innovations.

The bottom panel of Figure 1 compares continent-level innovation shares with population shares. The comparison is striking: innovation is far more concentrated than population. Europe and North America account for a large majority of global innovations in recent centuries, but their share of global

population is much smaller. Conversely, regions such as Asia and Africa have large population shares but contribute a comparatively small fraction of the innovations recorded in our dataset.

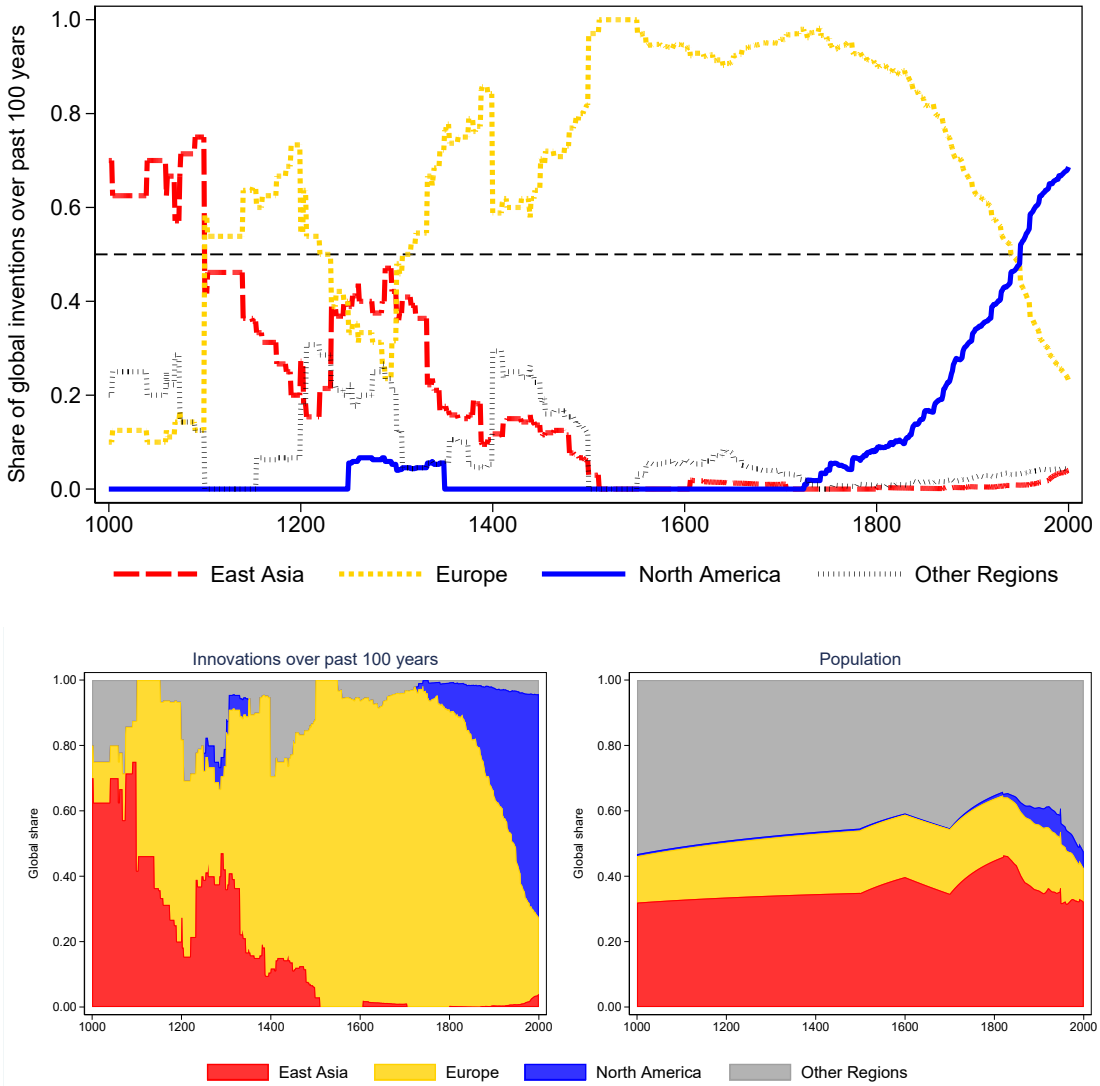
To formalize this comparison, Figure 2 plots Hirschman-Herfindahl indices (HHI) for innovation and population at the continent level. The HHI for innovation is high compared to that for population, especially since around 1500 AD. We also include the HHI for literature for additional information (as mentioned earlier, we use literature as an artistic example because this genre has a long history and has remained popular over time). Interestingly, literature is also spatially concentrated (more than population), but less so than technological innovation. This suggests that the extreme geographic concentration we document is not simply a feature of all intellectual or creative output but is particularly pronounced for technology.

Country level Figure 3 presents analogous evidence at the country level for the period 1500–2000, using cumulative innovations over the preceding 50 years. The top panel shows individual countries' shares of global innovations. Several patterns stand out. The United Kingdom was the leading innovator from the Industrial Revolution through the 19th century. Germany and France also contributed significant shares of innovations during this period. Subsequently, however, the United States rose steadily and became the single largest innovating nation by the early 20th century, and accounting for a dominant share of global innovations in the latter half of the century. The bottom panel of Figure 3 compares innovation shares with population and GDP shares for countries with consistent GDP data coverage. Innovation is more concentrated than both population and GDP. Even among high-income countries, innovation shares differ dramatically from GDP or population shares, indicating that the concentration of innovation is not merely a reflection of economic size.

Figure 4 presents HHI at the country level, reinforcing this observation. The innovation HHI is consistently higher than the population HHI across the sample period. As at the continent level, the HHI for literature lies between the innovation and population HHIs: the concentration of technological innovation exceeds that of artistic production.[2] Compared with Figure 2, the gap between the innovation and population HHIs is more pronounced at the country level than at the continent level. At the country level, there are many units, and population is spread broadly across them, which contributes to a low population HHI; meanwhile, innovation is dominated by a small number of countries, pushing the innovation HHI far above. At the continent level, aggregation compresses this variation: population becomes more concentrated (e.g., Asia alone accounts for a majority of the world's population), rais-

[2]Historical GDP data from Maddison is available only for subset of major countries. Therefore, we cannot compute the HHI for the full cross section covered by data on population, innovation and artistic production.

Figure 1. Innovation Concentration at the Continent Level

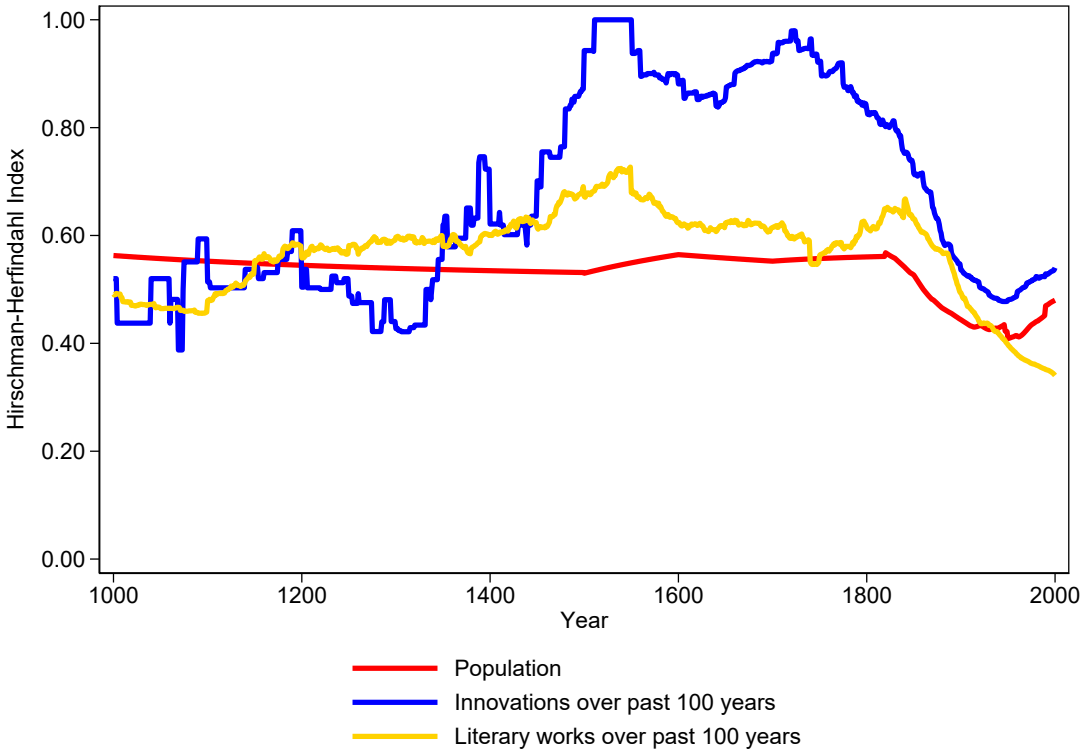


Notes: The figure illustrates the weight of continents in innovation and population between years 1000 and 2000. The top panel shows the share of continents in global innovations. The bottom panel shows stacked shares for global innovation and population shares. Global innovations are measured as cumulative sum over the past 100 years, to address lumpiness in the time series in early parts of the sample.

ing the population HHI; meanwhile, the concentration of innovation is diluted as within-continent heterogeneity—such as the gap between the United States and other countries in North America, or between leading and lagging innovators within Europe—is absorbed into continental aggregates.

U.S. state level The concentration patterns documented across countries also manifest within a single country. Figure IA5 presents innovation and population shares across U.S. states from 1800 to 2000, using cumulative innovations over the preceding 20 years. The top panel shows that a handful

Figure 2. Hirschman-Herfindahl Indices at the Continent Level

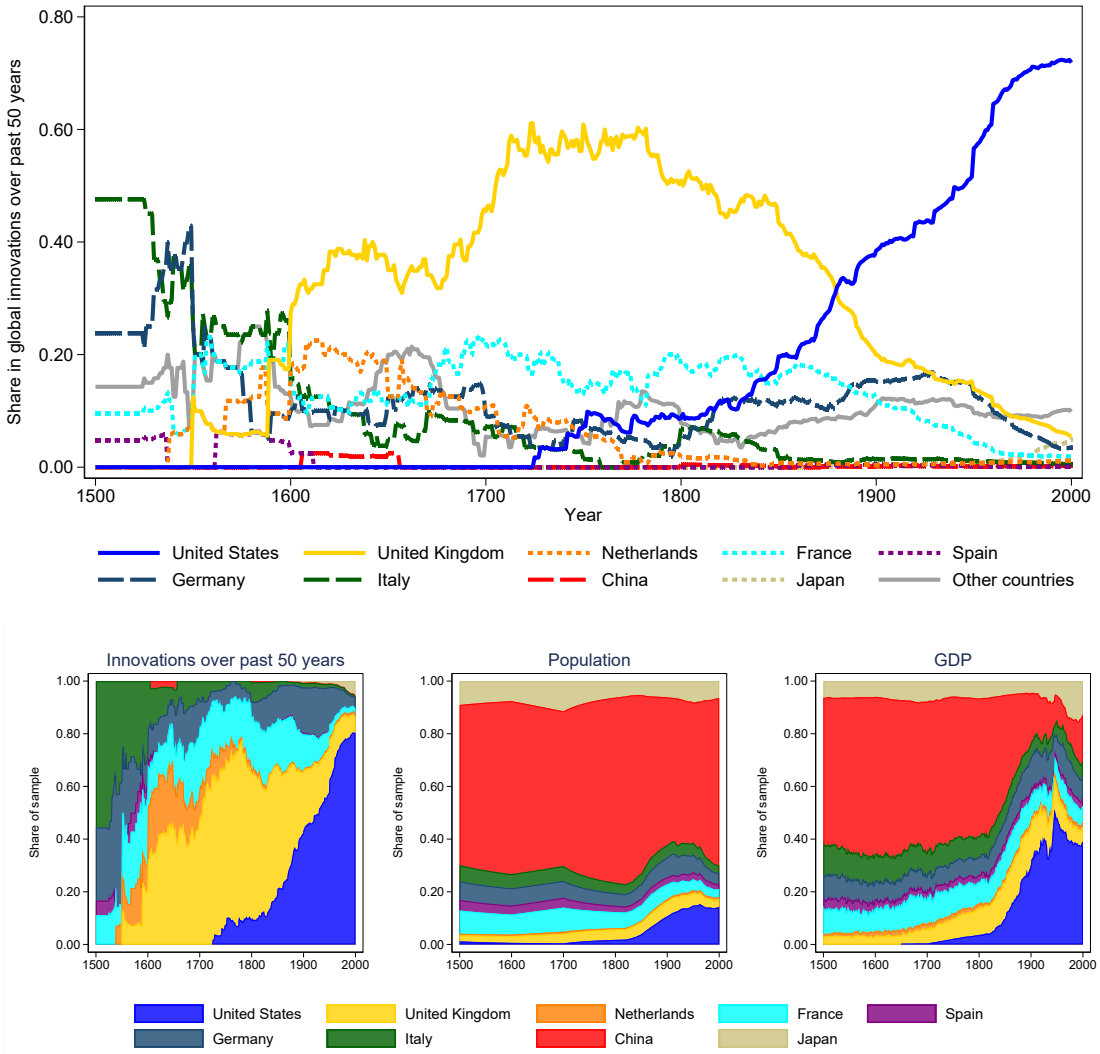


Notes: The figure plots Hirschman-Herfindahl concentration indices of innovation, population and literature at the continent level between years 1000 and 2000. The underlying sample of countries is restricted to those for which we have consistent population data for the full sample. Innovations and literature are measured as cumulative sum over the past 100 years, to address lumpiness in the time series in early parts of the sample.

of states account for the bulk of national innovations, with the leading states shifting over time. The bottom panel compares state-level innovation shares with population shares: as at the country and continent levels, innovation is substantially more concentrated than population. Figure IA6 confirms this with the HHI: the innovation HHI at the state level also lies above the population HHI throughout the sample.

Taken together, these findings show that innovation is one of the most spatially concentrated forms of economic activity. Moreover, the geographic center of innovation shifts over the span of centuries—from Asia in earlier periods, to Western Europe during the Scientific and Industrial Revolutions, and to the United States in the 20th century. This geographic dynamism suggests that the factors driving innovation concentration are themselves evolving over time.

Figure 3. Innovation Concentration at the Country Level



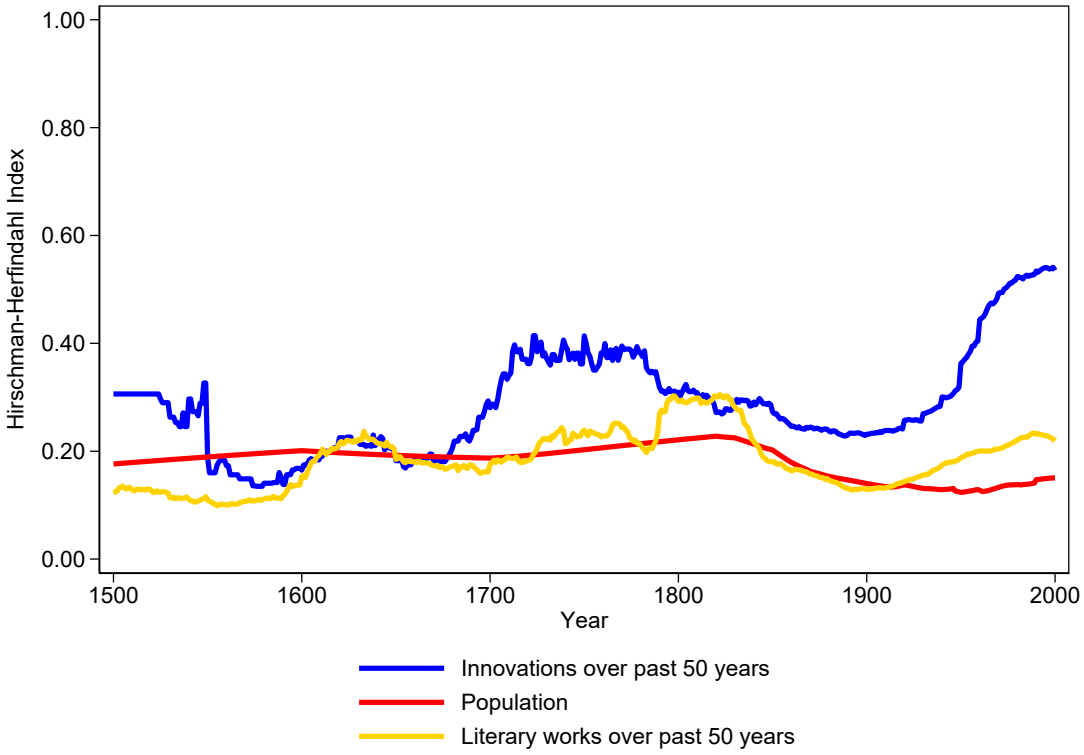
Notes: The figure illustrates the weight of different countries in innovation, population and GDP between 1500 and 2000. The top panel shows the share of countries in global innovations. The bottom panel stacked shares for selected countries with consistent coverage of GDP data for innovations, population and GDP. Global innovations are measured as cumulative sum over the past 50 years, to address lumpiness in the time series in early parts of the sample.

3.2 Innovations per Capita over Time

Results in Section 3.1 reveal that the total volume of innovation is highly concentrated across space. In this section, we examine innovations per capita and its evolution over time. This measure captures a region's or country's innovation intensity, which also points to striking patterns of rise and decline.

Continent level Figure 5 plots innovations per capita by continent from 1000 to 2000 AD, using cumulative innovations over the preceding 100 years relative to population. Europe had the highest

Figure 4. Hirschman-Herfindahl Indices at the Country Level

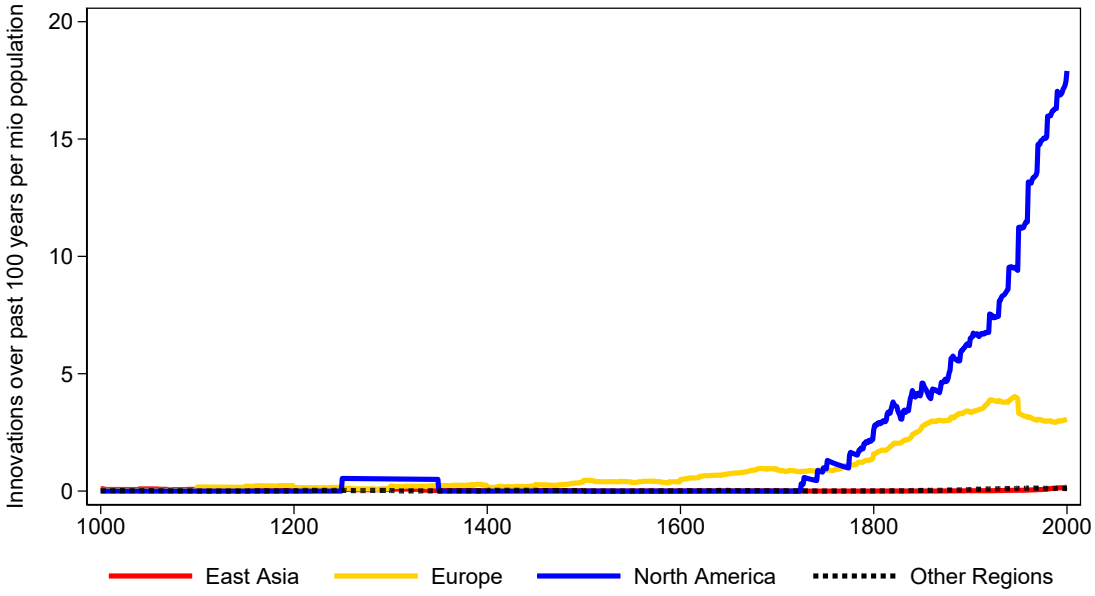


Notes: The figure plots Hirschman-Herfindahl concentration indices of innovation, population and literature at the country level between years 1500 and 2000. The sample of countries is restricted to those for which we have consistent population data for the full sample. Innovations and literature are measured as cumulative sum over the past 50 years, to address lumpiness in the time series in early parts of the sample.

innovations per capita for centuries, reflecting its central role in the Scientific Revolution and the successive waves of the Industrial Revolution. In the 20th century, however, North America surpassed Europe in innovations per capita, driven by the rapid ascent of the United States. Other continents have had substantially lower innovations per capita throughout the sample period.

Country level Figure 6 shows innovations per capita at the country level from 1500 to 2000. This figure reveals two of the most striking facts in our data. First, innovations per capita has decreased relative to one century ago in many European countries. The United Kingdom, for example, had extraordinarily high innovations per capita during and after the Industrial Revolution, but its innovation intensity has declined markedly since then. This decline is remarkable, as European countries did not necessarily experience drastic decay of institutional quality. Moreover, if Wikipedia titles have better coverage for recent time periods, this tendency should make innovations per capita in our dataset higher in recent time periods. Second, the United States has experienced a steady increase in innovations per

Figure 5. Innovations per Capita by Continent

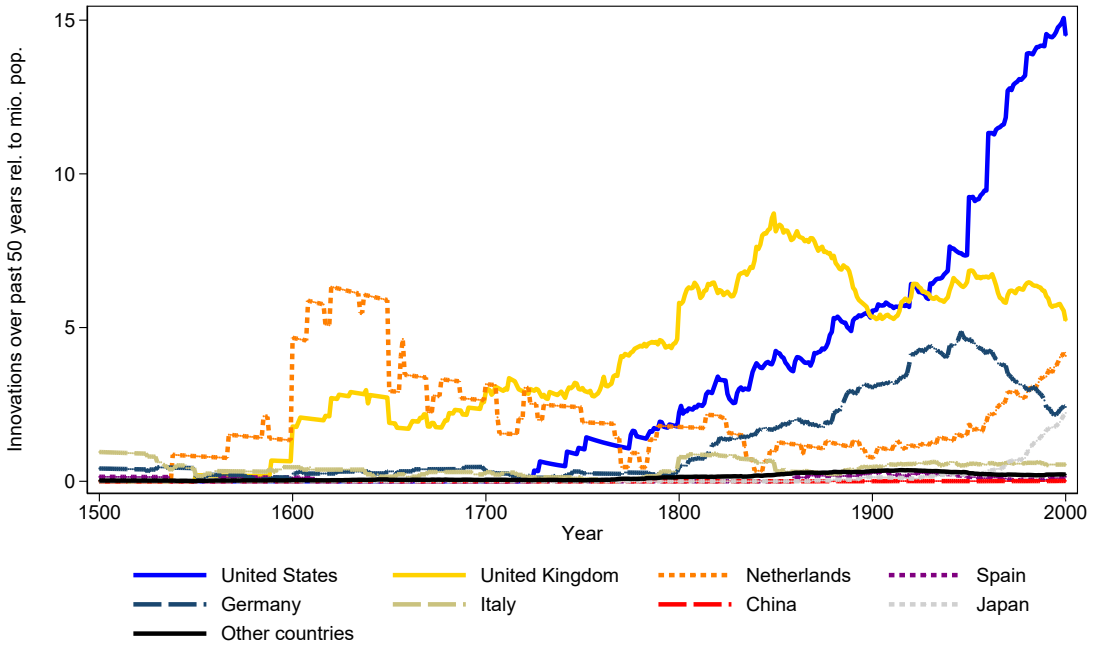


Notes: The figure plots the number of innovations conceived in a given region relative to that region's population between years 1000 and 2000. Innovations are measured as cumulative sum over the past 100 years, to address lumpiness in the time series in early parts of the sample.

capita since the founding of the country, a trajectory that has been quite persistent over more than two centuries. The vast heterogeneity in innovations per capita across countries is noteworthy, and it persists even among high-income countries.

U.S. state level Figure IA7 shows innovations per capita across U.S. states from 1800 to 2000. Large differences in innovations per capita emerge across states, mirroring the cross-country patterns. The timing of different states' innovation trajectories aligns closely with major historical developments. In the early 19th century, northeastern states such as Massachusetts and Connecticut—the cradle of American industrialization—had the highest innovations per capita. As the country expanded westward and new industries emerged, midwestern and mid-Atlantic states rose in prominence. In the 20th century, California experienced a dramatic rise in innovations per capita, coinciding with the growth of the aerospace industry, Silicon Valley, and the broader technology sector. New Mexico also stands out with a sharp increase in innovations per capita in the mid-20th century, reflecting the establishment of Los Alamos National Laboratory during the Manhattan Project and Sandia National Laboratories—an example of government-driven innovation transforming a state's innovative profile. These shifts across states echo the broader pattern observed across countries: the geographic center of innovation is dynamic, responding to evolving economic opportunities and the accumulation of local innovative

Figure 6. Innovations per Capita by Country



Notes: The figure plots the number of innovations conceived in a given country relative to that country's population between years 1500 and 2000. Innovations are measured as cumulative sum over the past 50 years, to address lumpiness in the time series in early parts of the sample.

capacity.

3.3 Spatial Correlation of Innovation over Time

We next examine the spatial correlation of innovation activity—that is, whether countries in geographic proximity tend to experience correlated changes in innovation over time. In principle, two opposing forces shape the spatial correlation of innovation. On the one hand, agglomeration forces such as knowledge spillovers and labor market pooling generate positive spatial correlation: innovation in one location spills over to nearby locations, causing proximate countries to innovate together. On the other hand, if innovation draws on a mobile factor—such as financial or human capital that flows toward regions with favorable conditions—then a surge in innovation in one country may come at the expense of its neighbors, generating negative spatial correlation. The sign and evolution of the spatial correlation can therefore shed light on which force predominates.

To investigate this, we estimate a regression relating the 20-year growth in innovation in country i to the 20-year growth in innovation in another country j, interacted with the geographic distance

between them:

$$\Delta_{20} \log(x_{i,t}) = \beta_1 \Delta_{20} \log(x_{j,t}) + \beta_2 \Big(\Delta_{20} \log(x_{j,t}) \times d_{i,j} \Big) + \boldsymbol{\gamma} \boldsymbol{z}_{i,j,t} + \epsilon_{i,j,t}, \tag{1}$$

where $\Delta_{20} \log(x_{i,t})$ denotes the 20-year change in the logarithm of cumulative innovations in country i, $d_{i,j}$ is the geographic distance between the capitals of countries i and j, and $\boldsymbol{z}_{i,j,t}$ captures additional controls. The data are observed at 20-year intervals. The estimated spatial elasticity $\hat{\beta}_1 + \hat{\beta}_2 d$ captures how the responsiveness of innovation growth in one country to innovation growth in another varies with distance.

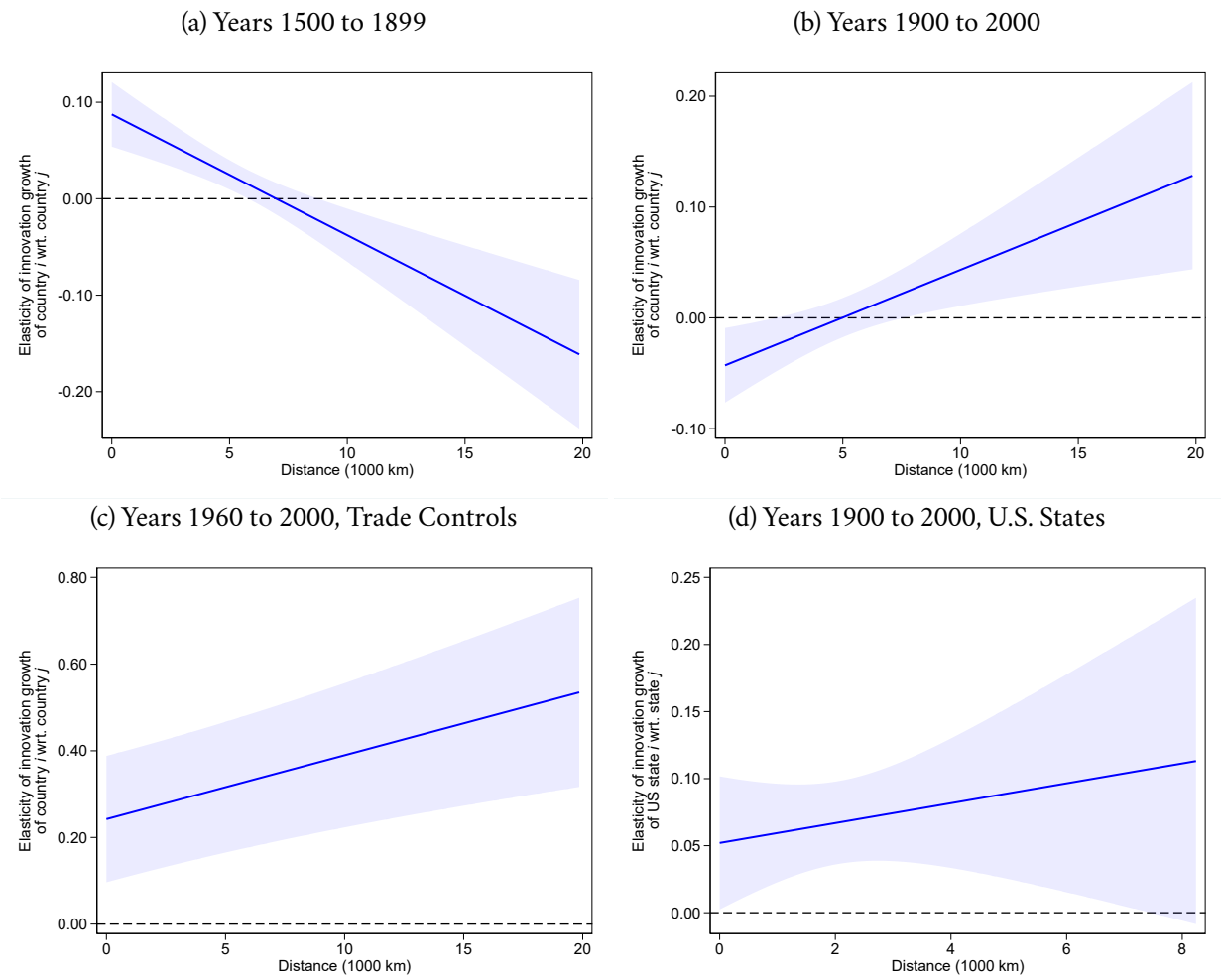
Figure 7 presents the estimated spatial elasticity as a function of distance, along with 95% confidence intervals based on robust standard errors. Panel (a) shows results for the historical period 1500–1899. Innovation growth in one country is positively associated with innovation growth in nearby countries, and this association declines with geographic distance. This pattern is consistent with agglomeration forces and knowledge spillovers dominating during this period.

Panel (b) presents analogous estimates for the modern period 1900–2000, where a striking change emerges: the spatial elasticity turns negative for nearby countries. That is, when a country experiences faster innovation growth, its geographic neighbors tend to experience slower innovation growth. This sign reversal is consistent with a mobile factor increasingly dominating agglomeration effects in the modern era. As innovation accelerates in one country, mobile resources—such as financial capital or skilled inventors—may flow toward that country and away from its neighbors, generating a negative spatial correlation at short distances.

Panel (c) addresses the concern that the spatial correlation may reflect factors correlated with distance rather than geographic proximity per se, such as trade linkages, cultural affinity, or institutional similarities. For the period 1960–2000, where trade-related data are available, we add controls for interactions of innovation growth with common language, common currency, colonial relationships, shared religion, and bilateral trade intensity. The negative spatial correlation for nearby countries persists after the inclusion of these controls, suggesting that the mobile factor channel operates beyond these observable dimensions.

Finally, Panel (d) examines spatial correlation within the United States across states for 1900–2000. A similar pattern emerges: innovation growth in one state is negatively associated with innovation growth in nearby states, consistent with mobile factors flowing toward innovating states within the country.

Figure 7. Innovation Comovement



Notes: The figure visualizes interaction effects estimated using the following regression: $\Delta_{20}\log(x_{i,t}) = \beta_1\Delta_{20}\log(x_{j,t}) + \beta_2\Big(\Delta_{20}\log(x_{j,t})\times d_{i,j}\Big) + \boldsymbol{\gamma}\boldsymbol{z}_{i,j,t} + \epsilon_{i,j,t}$, where $\Delta_{20}\log(x_{i,t})$ denotes the 20-year change in the logarithm of innovations over the past 20 years in country i at year t. Analogously, $\Delta_{20}\log(x_{j,t})$ denotes the same measure for another country $j \neq i$ and $d_{i,j}$ captures the geographical distance between both countries, specifically the coordinates of their modern day capitals. Any other time-pair-specific controls are captured by $\boldsymbol{z}_{i,j,t}$. The frequency t of the data is 20 year intervals. Each panel visualizes the estimated spatial elasticity of innovation growth $\hat{\beta}_1 + \hat{\beta}_2 d$ for different distances d alongside its 95% confidence interval based on robust standard errors. Panel (a) shows estimates obtained on the historical sample covering years between 1500 and 1899 where $\boldsymbol{z}_{i,j,t}$ includes a constant. Panel (b) shows estimates obtained on the modern sample covering years between 1900 and 2000 where $\boldsymbol{z}_{i,j,t}$ includes a constant. Panel (c) shows estimates obtained for years between 1960 and 2000 for which trade-related controls are available. Specifically, $\boldsymbol{z}_{i,j,t}$ includes a constant and interactions of $\Delta_{20}\log(x_{j,t})$ with i) a dummy indicating that the country pair shares a common language, ii) a dummy indicating that the country pairs shares a common currency, iii) a dummy indicating that the country pair is or was in a colonial relationship, iv) a dummy indicating that the country pair shares the same main religion, and v) a measure of bilateral trade flow intensity from trade literature $\log((f_{i,j}f_{j,i})/(\text{GDP}_i\text{GDP}_j))$. Panel (d) repeats estimation underlying panel (b) for U.S. state level data.

The shift from positive to negative spatial correlation over time is a revealing pattern. It suggests that in the historical period, when capital and labor were less mobile, agglomeration spillovers dominated and innovation tended to cluster and spread locally. In the modern era, as factor mobility increased, the competition for mobile resources became the dominant force, such that innovation gains in one location come partly at the expense of nearby locations. This evidence is consistent with the existence of an elastic, mobile factor—a theme we return to in the next section.

3.4 Heterogeneity in Different Types of Innovation

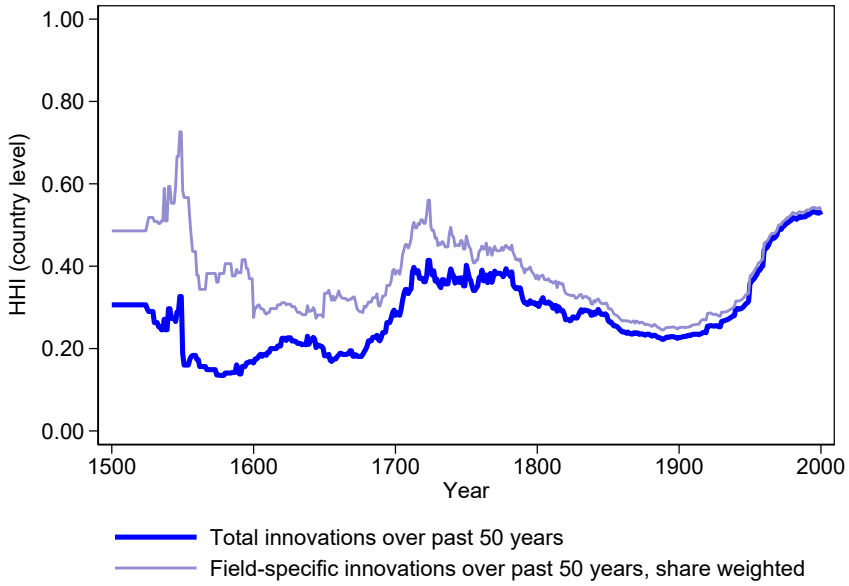
Finally, we investigate whether the spatial concentration patterns documented above differ across types of innovation. We consider two dimensions of heterogeneity: private versus public innovations, and the technological field or domain of application.

Private versus public innovations A natural question is whether the spatial concentration patterns documented above are driven by private or public sector innovation—and whether the two follow different spatial dynamics. We define private innovations as those attributed to companies and individual inventors, which respond more directly to market incentives and local economic conditions; and public innovations as those by universities, government entities, or pertaining to military and defense or basic research.

Figure IA12 shows continent-level innovation shares separately for the two types. The broad patterns are strikingly similar: both private and public innovations exhibit the same shifting centers of gravity documented for total innovations. A notable difference emerges in timing, however—dominance in private innovation activity tends to lead dominance in public innovation. Figures IA13 and IA14 examine the dynamic relationship between private and public innovation within countries. Both types exhibit high persistence over multi-decade horizons, but the cross-correlograms reveal that private innovation is a positive predictor of subsequent public innovation, with the reverse pattern less pronounced. In episodes where countries rise to global innovation leadership, private innovation increases more sharply and is more tightly timed around the transition. These patterns are inconsistent with the view that private innovation is a mere byproduct of public research—if anything, government and university activity has tended to accompany rather than precede surges in private invention.

Concentration across innovation fields A related question is whether the overall concentration of innovation is driven by concentration in a few specific technological fields, or whether it reflects a broad pattern across many domains. For instance, if the dominant innovation fields at a given point in time

Figure 8. HHI for Total and Field Innovations



Notes: The figure plots Hirschman-Herfindahl concentration indices of total innovation and weighted field-specific innovations at the country level between years 1500 and 2000. The HHI for field-specific innovations is computed as the weighted average of field-specific innovation HHIs where yearly weights are the share of field innovations in total innovations. Innovations are measured as cumulative sum over the past 50 years, to address lumpiness in the time series in early parts of the sample.

happen to be concentrated in a few countries—say, information technology in the United States—the overall HHI could be high even if each country's share within individual fields were relatively even. To investigate this, Figure 8 compares the country-level HHI for total innovation with the weighted average of field-specific HHIs, where the weights are the share of each field in total innovations in each year. If the overall concentration were driven by composition effects across fields, the weighted field-specific HHI would fall well below the total HHI. The figure shows that the field-specific HHI tracks the total HHI closely, indicating that the concentration of innovation is broad-based across technological domains rather than being an artifact of the dominance of any particular field.

Decomposing global and local factors To further assess the role of field composition, we decompose innovation activity into a global domain-specific component and a local country-specific component in Appendix Section IA3.2. Specifically, we estimate a fixed effects regression of the logarithm of innovations in domain d in country c at year t:

$$\log(i_{c,t,d}) = \underbrace{\gamma_{d,t}}_{\text{global domain}} + \underbrace{\gamma_{c,t}}_{\text{local country}} + e_{c,t,d}, \tag{2}$$

which yields an R^2 of 0.88. The implied variance decomposition of innovation growth shows that the local country component accounts for a substantially larger share of the variance than the global domain component. In other words, the rise and fall of innovation in a given country is driven primarily by country-specific factors rather than by the country happening to specialize in globally ascendant fields. Counterfactual exercises reinforce this observation: muting the global domain component leaves country-level innovation trajectories largely unchanged, whereas muting the local country component dramatically alters them. Global domain-specific shocks matter quantitatively only for small countries. This decomposition complements the field-level HHI analysis and confirms that the spatial concentration of innovation is fundamentally a local, country-specific phenomenon.

3.5 Additional Results

The preceding sections document stark patterns in the spatial distribution of innovation: high geographic concentration, shifting centers of innovation, and—in the modern era—negative spatial correlation consistent with mobile factors. Two natural questions follow. First, what observable country characteristics are associated with higher innovation intensity? Identifying the correlates of innovations per capita can shed light on which factors—among economic development, human capital, and different forms of financial development—are most closely linked to a country's innovative output. Second, how much does innovation matter for subsequent economic performance? If the innovations captured by our Wikipedia-based measure are economically meaningful, they should have predictive power for future output growth.

Country characteristics and innovation intensity We explore the first question by regressing innovations per capita on social, economic, and financial development indicators for 17 advanced economies over the period 1870–2000 (Appendix Table IA1). All variables are standardized, so coefficient magnitudes are directly comparable across regressors. This exercise is not intended to establish causal relationships, but rather to identify which country characteristics most strongly co-move with innovation intensity—and, in particular, whether the data point toward a specific mobile factor.

Population size and real GDP per capita have limited explanatory power in most specifications. Two variables stand out. First, universities per capita is positively and significantly associated with innovation intensity in the cross-sectional specifications, though the effect weakens within countries over time—suggesting that the cross-country variation in higher education institutions is more informative than changes within a given country. Second, and perhaps most intriguingly, stock market capitalization

relative to GDP is a consistently positive correlate of innovation intensity, with significant coefficients both across and within countries. By contrast, bank lending relative to GDP shows no significant association with innovations per capita in any specification. The stronger role of equity markets relative to bank lending is consistent with the idea that risk-tolerant financial capital—channeled through stock markets rather than bank loans—may be a particularly important mobile factor for innovation. These patterns complement the spatial correlation evidence in Section 3.3 and point to financial capital as a plausible candidate for the elastic, mobile factor behind the geographic concentration of innovation.

Innovation and subsequent output growth We turn next to the relationship between innovation and future economic performance. If our Wikipedia-based measure captures economically meaningful innovations, countries with more innovative activity should experience faster subsequent growth—even after controlling for current income levels and recent growth trajectories. In the Internet Appendix, we examine this using local projections that relate a country's future real GDP per capita growth to its recent innovation activity, controlling for country and year fixed effects, contemporaneous and lagged GDP growth, and the level of income to account for catch-up effects (Appendix Figure IA2). The sample covers 1500 to 2000, with GDP data from Maddison.

The results confirm this prediction: countries with more innovations over the past decade experience higher subsequent growth in real GDP per capita, with the effect building over longer horizons. When we decompose innovations by type, an interesting asymmetry emerges. Private innovations—those by companies and individual inventors—show a stronger predictive relationship with future growth than public innovations, which include those by universities, government entities, or pertaining to military and basic research. This pattern is consistent with the view that privately driven innovations are more closely tied to productive capacity and market opportunities, while public innovations may serve broader purposes—such as national defense or fundamental scientific knowledge—that do not translate as directly into measured output growth. This finding also validates our data: the Wikipedia-based innovation measure captures economically meaningful variation in innovative activity, beyond what is already reflected in contemporaneous income levels.

If innovations drive growth, a natural question is whether this benefit accrues mainly to the places where innovations originate, or whether it spreads more broadly through diffusion. As documented in Section 3.1, innovation is far more concentrated than GDP: a handful of countries account for the bulk of global innovations but a much smaller share of global output, and the innovation HHI persistently exceeds the population and GDP HHIs at the country level (Figure 4). Diffusion of ideas across borders could help account for this wedge.

We examine the patterns of diffusion in Appendix Figure IA3. We observe that the global innovation leader's activity predicts GDP per capita growth in *other* countries, controlling for each country's own innovation and income level. In the pre-1900 sample, the relationship is modest. In the 1900–2000 sample, it is substantially larger: an increase in the leader's innovations per capita is associated with meaningfully higher growth abroad, with the effect building over longer horizons. Meanwhile, Appendix Figure IA4 shows that a country's own innovations are a stronger predictor of its subsequent growth before 1900 than after. In the modern era, the relationship between own innovation and future local growth attenuates considerably. These two findings suggest that the economic gains from innovation have become less tied to their place of origin over time, consistent with accelerating diffusion. These patterns help explain two features of the concentration evidence. First, the spatial concentration of output is lower than the concentration of innovation (Figure 4), because ideas spread—output reflects not only local invention but also the adoption of technologies from elsewhere. Second, the gap between innovation and output concentration at the country level can widen over time: as diffusion becomes faster, the economic benefits of innovation disperse more broadly even as their origins remain geographically concentrated.

4 A Framework with Agglomeration and Diffusion

There is a finite set of locations, and each location ℓ is endowed with a measure $L_{\ell t}$ of individuals that is growing at rate γ. The individuals can choose among three tasks. Let $L_{\ell t}^R$ be the measure that choose to conduct research, $L_{\ell t}^A$ be the measure that choose to adopting varieties, and $L_{\ell t}^Y$ be the measure that choose to work in production, so that

$$L_{\ell t}^Y + L_{\ell t}^R + L_{\ell t}^A = L_{\ell t}$$

There is a unit continuum of domains. Research generates inventions, which are techniques for particular domains. Each technique has a productivity, q. Let $F_{\ell t}(q)$ be the fraction of domains in location ℓ at time t for which the best technique has productivity weakly less than q.

Research creates new techniques. When a new technique arrives, its domain is randomly assigned and its productivity is a combination of an original component o and an insight q' drawn randomly from a source distribution with CDF $G_{\ell t}(q')$ according to $q = (q')^{\beta}\, o$. β represents the strength of diffusion, as it indexes the importance of insights from others in the productivity of new techniques.

New ideas where the original component is larger than o is

$$\alpha_{\ell t} o^{-\theta}$$

where

$$\alpha_{\ell t} \equiv L_{\ell t}^{R} \left(L_{\ell t-1}^{R} \right)^{\zeta} Z_{\ell t}^{R}$$

As shown in Buera and Oberfield (2020), if initial conditions are such that the frontier follows a Frechet distribution, $F_{\ell 0}(q) = e^{-\lambda_{\ell 0} q^{-\theta}}$, then the Frontier will remain Frechet, with

$$F_{\ell t}(q) = e^{-\lambda_{\ell t} q^{-\theta}}$$

where $\lambda_{\ell t}$ is the local stock of knowledge, which evolves according to

$$\lambda_{\ell t} - \lambda_{\ell t-1} = \alpha_{\ell t} \int_0^{\infty} x^{\beta\theta} dG_{\ell t}(x)$$

One simple theory for the source distribution $G_{\ell t}$ is that insights from each country $\tilde{\ell}$ are discounted by $\kappa_{\ell\tilde{\ell}} \geq 1$, giving

$$G_{\ell t}(q) = e^{-\left(\sum_{\tilde{\ell}} \lambda_{\tilde{\ell} t-1} \kappa_{\ell\tilde{\ell},t}^{-\theta}\right) q^{-\theta}}$$

[3] Adaptors (also called entrepreneurs) can access ideas from any location. When entrepreneurs pay a cost to develop a variety, they draw a random domain for that variety. Then they can access techniques for that domain from anywhere in the world, although techniques from abroad are less productive. If the best idea in location $\tilde{\ell}$ for domain j is $q_{j\tilde{\ell}t}$, then the entrepreneur in ℓ who has a variety that uses domain j is

$$\max_{\tilde{\ell}} \left\{ \frac{q_{j\tilde{\ell}t}}{\tau_{\tilde{\ell}\ell}} \right\}$$

Thus the fraction of domains for which the best technique available to entrepreneurs in ℓ has productivity weakly less than q is

$$H_{\ell t}(q) = e^{-\left(\sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t} \tau_{\ell\tilde{\ell},t}^{-\theta}\right) q^{-\theta}}$$

[3] Another theory, closer to Eaton and Kortum (1999) is that insights generated in $\tilde{\ell}$ become accessible in ℓ stochastically, with probability $\varphi_{\ell\tilde{\ell}}$ each period, so that

$$e^{-\left(\sum_{\tilde{\ell}} \sum_{s=1}^{\infty} (1-(1-\varphi_{\ell\tilde{\ell}})^s)\left(\lambda_{\tilde{\ell} t-s} - \lambda_{\tilde{\ell} t-s-1}\right)\right) q^{-\theta}}$$

For an entrepreneur in ℓ, the cost of creating a new variety is $\frac{1}{Z_{\ell t}^A}$. After paying this cost, the adopter draws a random domain from which to select a technique with which to produce the variety. The measure of varieties created in ℓ is

$$N_{\ell t} = Z_{\ell t}^A L_{\ell t}^A$$

The ability to produce a variety depreciates after a single period.

Each individual has Dixit-Stiglitz preferences across all varieties $Y_{\ell t} = \left(\int_0^{N_{\ell t}} y_{j\ell t}^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}$, so it faces price index $P_{\ell t} = \left(\int_0^{N_{\ell t}} p_{j\ell t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}$. Output is

$$Y_{\ell t} = N_{\ell t}^{\frac{1}{\varepsilon-1}} Q_{\ell t} L_{\ell t}^Y = \left(Z_{\ell t}^A L_{\ell t}^A \right)^{\frac{1}{\varepsilon-1}} Q_{\ell t} L_{\ell t}^Y$$

where $Q_{\ell t}$ is

$$Q_{\ell t} \equiv \left[\int_0^\infty q^{\varepsilon-1} dH_{\ell t}(q) \right]^{\frac{1}{\varepsilon-1}} = \Gamma \left(1 - \frac{\varepsilon-1}{\theta} \right)^{\frac{1}{\varepsilon-1}} \left(\sum_{\tilde{\ell}} \lambda_{\tilde{\ell} t} \tau_{\ell\tilde{\ell},t}^{-\theta} \right)^{\frac{1}{\theta}}$$

4.1 Market Structure

Let $w_{\ell t}^R$, $w_{\ell t}^Y$ and $w_{\ell t}^A$ be the local wages of each of the three activities.

For each variety, the adaptor is a monopolist with productivity $q_{j\ell t}$ drawn from the distribution with CDF $H_{\ell t}$. The adaptor faces demand is $Y_{\ell t} P_{\ell t}^{\varepsilon} p_{j\ell t}^{-\varepsilon}$ and unit cost is $\frac{w_{\ell t}^Y}{q_{j\ell t}}$, so its price is $p_{j\ell t} \mu_{\ell t} \frac{w_{\ell t}^Y}{q_{j\ell t}}$, where $\mu_{\ell t}$ is the local markup (which may be smaller than the desired markup of $\frac{\varepsilon}{\varepsilon-1}$ if intellectual property protection is not so strong and a competitive fringe can create a substitute at a slightly higher cost). The adaptor's revenue net of cost is

$$Y_{\ell t} P_{\ell t}^{\varepsilon} p_{j\ell t}^{-\varepsilon} \left(p_{j\ell t} - \frac{w_{\ell t}^Y}{q_{j\ell t}} \right) = \mu_{\ell t}^{-\varepsilon} \left(\mu_{\ell t} - 1 \right) Y_{\ell t} P_{\ell t}^{\varepsilon} \left(\frac{w_{\ell t}^Y}{q_{j\ell t}} \right)^{1-\varepsilon}$$

The price level is $P_{\ell t} = \left(\int_0^{N_{\ell t}} p_{j\ell t}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}} = \left(\int_0^{N_{\ell t}} \left(\mu_{\ell t} \frac{w_{\ell t}^Y}{q_{j\ell t}} \right)^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}} = \frac{N_{\ell t}^{\frac{1}{1-\varepsilon}} \mu_{\ell t} w_{\ell t}^Y}{Q_{\ell t}}$. Profit for an individual firm is

$$Y_{\ell t} P_{\ell t}^{\varepsilon} p_{j\ell t}^{-\varepsilon} \left(p_{j\ell t} - \frac{w_{\ell t}^Y}{a_{j\ell t}} \right) = Y_{\ell t} P_{\ell t}^{\varepsilon} \left(p_{j\ell t}^{1-\varepsilon} \right) \frac{\mu_{\ell t} - 1}{\mu_{\ell t}} = Y_{\ell t} P_{\ell t} P_{\ell t}^{\varepsilon-1} \left(\mu_{\ell t} \frac{w_{\ell t}^Y}{q_{j\ell t}} \right)^{1-\varepsilon} \frac{\mu_{\ell t} - 1}{\mu_{\ell t}}$$

For adaptors that have not yet drawn a domain, expected profit is

$$\frac{1}{N_{\ell t}}\int Y_{\ell t}P_{\ell t}^{\varepsilon}p_{\ell jt}^{1-\varepsilon}\frac{\mu_{\ell t}-1}{\mu_{\ell t}}dj=\frac{1}{N_{\ell t}}Y_{\ell t}P_{\ell t}\frac{\mu_{\ell t}-1}{\mu_{\ell t}}$$

and total expenditure on production labor is

$$Y_{\ell t}P_{\ell t}=\mu_{\ell t}w_{\ell t}^{Y}L_{\ell t}^{Y}$$

Adaptors may pay royalties to researchers who invented the techniques that they use. Let $T_{\ell'\ell t}$ be the total royalty payments from ℓ to ℓ' at t. So if entry cost is $\frac{w_{\ell t}^{A}}{Z_{\ell t}^{A}}$, then free entry implies

$$\frac{w_{\ell t}^{A}}{Z_{\ell t}^{A}}N_{\ell t}=\left(\mu_{\ell t}-1\right)w_{\ell t}^{Y}L_{\ell t}^{Y}-\sum_{\ell'}T_{\ell'\ell t}$$

and market clearing for entrepreneurship is

$$Z_{\ell t}^{A}L_{\ell t}^{A}=N_{\ell t}$$

which together yield

$$w_{\ell t}^{A}L_{\ell t}^{A}=\left(\mu_{\ell t}-1\right)w_{\ell t}^{Y}L_{\ell t}^{Y}-\sum_{\ell'}T_{\ell'\ell t}$$

There is free entry into research and ideas arrive at rate $\alpha_{\ell t}$, The payoff to all researchers in ℓ at t is

$$w_{\ell t}^{R}L_{\ell t}^{R}=\sum_{\ell'}T_{\ell\ell'}$$

4.2 Royalties

Suppose that when an entrepreneur in ℓ_0 uses a technique invented by a researcher in ℓ at t (i.e., within the period), the entrepreneur keeps a share $1-\eta_{\ell\ell_0 t}$ of the difference in profit between the best and second best techniques for that domain she uses, and pays the remainder to researcher. If the technique was invented before the current period, no royalties are paid. Building on Bernard et al. (2003), we show in Appendix IA4 total royalty payments from country ℓ_0 paid to researchers in ℓ is

$$T_{\ell\ell_0 t}=\eta_{\ell\ell_0 t}Y_{\ell_0 t}P_{\ell_0 t}\frac{\varepsilon-1}{\theta}\frac{\mu_{\ell_0 t}-1}{\mu_{\ell_0 t}}\frac{\left(\lambda_{\ell t}-\lambda_{\ell t-1}\right)\tau_{\ell\ell_0}^{-\theta}}{\sum_{\tilde{\ell}}\lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}$$

4.3 Labor Supply

Each individual chooses an activity, $n \in \{R, A, Y\}$. Individuals have heterogeneous tastes for the various activities. The benefit of working in activity n is the wage, $w_{\ell t}^n$, a systematic component of taste for the activity in the location, $B_{\ell t}^n$, and an idiosyncratic taste shock, ε_i^n. Each individual chooses the activity that maximizes welfare: activity

$$\max_n \log \frac{w_{\ell t}^n}{P_{\ell t}} + \log B_{\ell t}^n + \frac{1}{\sigma}\varepsilon_i^n$$

We assume that the idiosyncratic taste shocks are IID and follow a standard Gumbel distribution so that $\Pr\left(\varepsilon_i^n \leq \varepsilon\right) = e^{-e^{-\varepsilon}}$. Then $\Pr\left(n\right) =$. As a result, the probability of working in activity n is

$$\frac{\left(w_{\ell t}^n B_{\ell t}^n\right)^\sigma}{\sum_n \left(w_{\ell t}^{n'} B_{\ell t}^{n'}\right)^\sigma}$$

This implies

$$\begin{aligned}
\frac{L_{\ell t}^R}{L_{\ell t}^Y} &= \frac{\left(w_{\ell t}^R B_{\ell t}^R\right)^\sigma}{\left(w_{\ell t}^Y B_{\ell t}^Y\right)^\sigma} \\
\frac{L_{\ell t}^A}{L_{\ell t}^Y} &= \frac{\left(w_{\ell t}^A B_{\ell t}^A\right)^\sigma}{\left(w_{\ell t}^Y B_{\ell t}^Y\right)^\sigma}
\end{aligned}$$

or

$$\begin{aligned}
\left(\frac{L_{\ell t}^R}{L_{\ell t}^Y}\right)^{\frac{\sigma+1}{\sigma}} &= \frac{B_{\ell t}^R}{B_{\ell t}^Y}\frac{w_{\ell t}^R L_{\ell t}^R}{w_{\ell t}^Y L_{\ell t}^Y} \\
\left(\frac{L_{\ell t}^A}{L_{\ell t}^Y}\right)^{\frac{\sigma+1}{\sigma}} &= \frac{B_{\ell t}^A}{B_{\ell t}^Y}\frac{w_{\ell t}^A L_{\ell t}^A}{w_{\ell t}^Y L_{\ell t}^Y}
\end{aligned}$$

4.4 Only Local Royalties: Labor Demand

We now focus on an economy in which adaptors only pay royalties to local researchers. In that case,

$$T_{\ell t} = \eta_{\ell t}\frac{\varepsilon - 1}{\theta}\frac{\left(\lambda_{\ell t} - \lambda_{\ell t-1}\right)}{\sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}\left(\mu_{\ell t} - 1\right) w_{\ell t}^Y L_{\ell t}^Y$$

To simplify the next derivations, define $\phi_{\ell t} \equiv \eta_{\ell t}\frac{\varepsilon-1}{\theta}\frac{\left(\lambda_{\ell t} - \lambda_{\ell t-1}\right)}{\sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}$, so that $T_{\ell t} = \phi_{\ell t}\left(\mu_{\ell t} - 1\right) w_{\ell t}^Y L_{\ell t}^Y$

Total earnings of researchers and entrepreneurs are, respectively

$$w_{\ell t}^R L_{\ell t}^R = \phi_{\ell t} \left(\mu_{\ell t} - 1\right) w_{\ell t}^Y L_{\ell t}^Y$$

$$w_{\ell t}^A L_{\ell t}^A = \left(1 - \phi_{\ell t}\right) \left(\mu_{\ell t} - 1\right) w_{\ell t}^Y L_{\ell t}^Y$$

The labor supply and demand equations imply

$$\begin{aligned}
\left(\frac{L_{\ell t}^R}{L_{\ell t}^Y}\right)^{\frac{\sigma+1}{\sigma}} &= \frac{B_{\ell t}^R}{B_{\ell t}^Y} \frac{w_{\ell t}^R L_{\ell t}^R}{w_{\ell t}^Y L_{\ell t}^Y} = \frac{B_{\ell t}^R}{B_{\ell t}^Y} \phi_{\ell t} \left(\mu_{\ell t} - 1\right) \\
\left(\frac{L_{\ell t}^A}{L_{\ell t}^Y}\right)^{\frac{\sigma+1}{\sigma}} &= \frac{B_{\ell t}^A}{B_{\ell t}^Y} \frac{w_{\ell t}^A L_{\ell t}^A}{w_{\ell t}^Y L_{\ell t}^Y} = \frac{B_{\ell t}^A}{B_{\ell t}^Y} \left(1 - \phi_{\ell t}\right) \left(\mu_{\ell t} - 1\right)
\end{aligned}$$

These can be expressed as

$$\begin{aligned}
L_{\ell t}^R &= \left[\phi_{\ell t} \left(\mu_{\ell t} - 1\right) \frac{B_{\ell t}^R}{B_{\ell t}^Y}\right]^{\frac{\sigma}{1+\sigma}} L_{\ell t}^Y \\
L_{\ell t}^A &= \left[\left(1 - \phi_{\ell t}\right) \left(\mu_{\ell t} - 1\right) \frac{B_{\ell t}^A}{B_{\ell t}^Y}\right]^{\frac{\sigma}{1+\sigma}} L_{\ell t}^Y
\end{aligned}$$

These along with $L_{\ell t} = L_{\ell t}^R + L_{\ell t}^A + L_{\ell t}^Y$ give

$$\begin{aligned}
s_{\ell t}^R &\equiv \frac{L_{\ell t}^R}{L_{\ell t}} = \frac{\left[\phi_{\ell t} B_{\ell t}^R\right]^{\frac{\sigma}{1+\sigma}}}{\left[\phi_{\ell t} B_{\ell t}^R\right]^{\frac{\sigma}{1+\sigma}} + \left[\left(1 - \phi_{\ell t}\right) B_{\ell t}^A\right]^{\frac{\sigma}{1+\sigma}} + \left(\frac{1}{\mu_{\ell t} - 1} B_{\ell t}^Y\right)^{\frac{\sigma}{1+\sigma}}} \\
s_{\ell t}^A &\equiv \frac{L_{\ell t}^A}{L_{\ell t}} = \frac{\left[\left(1 - \phi_{\ell t}\right) B_{\ell t}^A\right]^{\frac{\sigma}{1+\sigma}}}{\left[\phi_{\ell t} B_{\ell t}^R\right]^{\frac{\sigma}{1+\sigma}} + \left[\left(1 - \phi_{\ell t}\right) B_{\ell t}^A\right]^{\frac{\sigma}{1+\sigma}} + \left(\frac{1}{\mu_{\ell t} - 1} B_{\ell t}^Y\right)^{\frac{\sigma}{1+\sigma}}} \\
s_{\ell t}^Y &\equiv \frac{L_{\ell t}^Y}{L_{\ell t}} = \frac{\left(\frac{1}{\mu_{\ell t} - 1} B_{\ell t}^Y\right)^{\frac{\sigma}{1+\sigma}}}{\left[\phi_{\ell t} B_{\ell t}^R\right]^{\frac{\sigma}{1+\sigma}} + \left[\left(1 - \phi_{\ell t}\right) B_{\ell t}^A\right]^{\frac{\sigma}{1+\sigma}} + \left(\frac{1}{\mu_{\ell t} - 1} B_{\ell t}^Y\right)^{\frac{\sigma}{1+\sigma}}}
\end{aligned}$$

Note that $s_{\ell t}^R$ is increasing in $\phi_{\ell t}$ (which is proportional to $\eta_{\ell t} \frac{\lambda_{\ell t} - \lambda_{\ell t-1}}{\sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t} \tau_{\ell\tilde{\ell},t}^{-\theta}}$), $\mu_{\ell t}$, and $B_{\ell t}^R$.

How does research evolve?

$$\begin{aligned}\lambda_{\ell t}-\lambda_{\ell t-1} &= \alpha_{\ell t}\int_0^\infty x^{\beta\theta}dG_{\ell t-1}(x)\\ &= L_{\ell t}^R\left(L_{\ell t-1}^R\right)^\zeta Z_{\ell t}^R\int_0^\infty x^{\beta\theta}dG_{\ell t-1}(x)\\ &= s_{\ell t}^R\left(s_{\ell t-1}^R\right)^\zeta L_{\ell t}L_{\ell t-1}^\zeta Z_{\ell t}^R\int_0^\infty x^{\beta\theta}dG_{\ell t-1}(x)\end{aligned}$$

This highlights the various forces that can impact the arrival rate of ideas.

- $Z_{\ell t}$ are institutional factors that affect the productivity of researchers
- $B_{\ell t}^R$ is modeled as a preference for research, stands in for cultural factors or norms that push people in to research
- $L_{\ell t}$ is scale
- $\int_0^\infty x^{\beta\theta}dG_{\ell t-1}(x)$ summarizes insights from others
- $\eta_{\ell t}$ indexes intellectual property protection
- $\mu_{\ell t}$ which indexes how much profit is available to entrepreneurs, which will translate into royalties for researchers
- $\frac{(\lambda_{\ell t}-\lambda_{\ell t-1})}{\sum_{\tilde{\ell}}\lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}$ summarizes discouragement from being far from frontier. If others are far ahead, it will be unlikely that the ideas of a researcher will actually be used.

Income per capita is

$$\frac{Y_{\ell t}}{L_{\ell t}}=\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{\frac{1}{\varepsilon-1}}\left(\sum_{\tilde{\ell}}\lambda_{\tilde{\ell}t}\tau_{\ell\tilde{\ell},t}^{-\theta}\right)^{\frac{1}{\theta}}\left(Z_{\ell t}^A s_{\ell t}^A\right)^{\frac{1}{\varepsilon-1}}s_{\ell t}^Y L_{\ell t}^{\frac{1}{\varepsilon-1}}$$

4.5 Detrending

The source distribution is

$$G_{\ell t}(q)=e^{-\left(\sum_{\tilde{\ell}}\lambda_{\tilde{\ell}t-1}\kappa_{\ell\tilde{\ell},t}^{-\theta}\right)q^{-\theta}}$$

Thus the contribution of insights is

$$\int_0^\infty x^{\beta\theta} dG_{\ell t}(x) = \Gamma(1-\beta)\left(\sum_{\tilde{\ell}} \lambda_{\tilde{\ell}t-1}\kappa_{\ell\tilde{\ell},t}^{-\theta}\right)^{\beta}$$

Suppose that each location's labor grows at some fixed rate γ so that $L_{\ell t} = (1+\gamma) L_{\ell t-1}$. Then we can define

$$\begin{aligned}
\hat{L}_{\ell t} &= \frac{L_{\ell t}}{(1+\gamma)^t} \\
\hat{\lambda}_{\ell t} &= \frac{\lambda_{\ell t}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}t}} \\
\hat{Y}_{\ell t} &= \frac{Y_{\ell t}}{(1+\gamma)^{\left(\frac{1+\zeta}{1-\beta}\frac{1}{\theta}+\frac{1}{\varepsilon-1}\right)t}}
\end{aligned}$$

Then we can express $\phi_{\ell t}$ as

$$\phi_{\ell t} = \eta_{\ell t}\frac{\varepsilon-1}{\theta}\frac{\hat{\lambda}_{\ell t} - \frac{1}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\hat{\lambda}_{\ell t-1}}{\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}$$

The expressions for $s_{\ell t}^n$ are unchanged. The law of motion for $\hat{\lambda}_{\ell t}$ becomes

$$\hat{\lambda}_{\ell t} - \frac{\hat{\lambda}_{\ell t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}} = s_{\ell t}^R\left(s_{\ell t-1}^R\right)^{\zeta}\hat{L}_{\ell t}\hat{L}_{\ell t-1}^{\zeta}(1+\gamma)^{-\zeta}Z_{\ell t}^R\Gamma(1-\beta)\left(\sum_{\tilde{\ell}}\frac{\hat{\lambda}_{\tilde{\ell}t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\kappa_{\ell\tilde{\ell},t}^{-\theta}\right)^{\beta}$$

and output per worker becomes

$$\hat{Y}_{\ell t} = \Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{\frac{1}{\varepsilon-1}}\left(\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell}t}\tau_{\ell\tilde{\ell},t}^{-\theta}\right)^{\frac{1}{\theta}}\left(\frac{s_{\ell t}^A}{\chi_{\ell t}}\right)^{\frac{1}{\varepsilon-1}} s_{\ell t}^Y\hat{L}_{\ell t}^{\frac{1}{\varepsilon-1}}$$

4.6 Convergence Among Countries and Continents

In this section we show that when countries within a continent are more closely connected with each other than countries across continents, then leadership in innovation is more persistent among continents than among countries. We show this formally using a simple symmetric example and log-linearizing around the steady state.

In the simple example, we suppose that the costs of distance remain constant over time and the $\tau_{\ell\tilde{\ell},t} = \tau_{\ell\tilde{\ell},t}$ and $\kappa_{\ell\tilde{\ell},t} = \kappa_{\ell\tilde{\ell},t}$, that entrepreneurs' markups are constant over time $\mu_{\ell t} = \mu_\ell$, that there

are no dynamic agglomeration affects, $\zeta = 0$, and that adopters can only access foreign technologies with a one period lag.

Country ℓ's innovation is measured by $x_{\ell,t} \equiv \tilde{\lambda}_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \tilde{\lambda}_{\ell t-1}$, where $\tilde{\lambda}$ is the log deviation of detrended knowledge $\hat{\lambda}_{\ell t}$ from steady state. In the Appendix, we show that once we log linearize, innovation in each country follows the process:

$$x_{\ell t} = -B_{1\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\ell,t-1} + B_{2\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell}t-1} + \varepsilon_{\ell t} \tag{3}$$

with

$$\varsigma_{\ell t} \equiv \Theta_{1\ell} \tilde{\eta}_{\ell t} + \Theta_{2\ell} \left\{ \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} \right\}$$

summarizes the disturbances that affect incentives to innovate: intellectual property protection, η_{ℓ} ; taste for innovation, $s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right)$; market size, $L_{\ell t}$; and research productivity, $Z_{\ell t}^{R}$. $B_{1\ell}$, $B_{2\ell}$, $\Theta_{1\ell}$ and $\Theta_{2\ell}$ are ℓ-specific constant that depend on its steady state values.

(3) shows two sources of spillovers across countries. The summarizes the impact of technology abroad on incentives to incentives to be an entrepreneur. When good technology is available abroad, the payoff to using that technology to create a variety is higher. This reduces incentives to innovate. The second term summarizes spillovers in the research process. If the state of technology is better, research is more productive because there are better insights to build on.

We next turn to the persistence of leadership in innovation. Taking quasidifferences of (3) gives

$$x_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} x_{\ell t-1} = -B_{1\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} x_{\ell,t-1} + B_{2\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} x_{\tilde{\ell}t-1} + \varepsilon_{\ell t}$$

with $\varepsilon_{\ell t} = \varsigma_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \varsigma_{\ell t}$. To make analytical progress, we focus on the special case in which there are N continents each with J countries. Suppose that costs of distance are such that $\tau_{\ell\tilde{\ell}} = \tau_{within}$ and $\kappa_{\ell\tilde{\ell}} = \kappa_{within}$ if ℓ and $\tilde{\ell}$ are on the same continent, and $\tau_{\ell\tilde{\ell}} = \tau_{across}$ and $\kappa_{\ell\tilde{\ell}} = \kappa_{across}$ if ℓ and $\tilde{\ell}$ are on different continents. The countries are otherwise symmetric, so that $B_{1\ell} = B_1$ and $B_{2\ell} = B_2$. The system can be expressed as

$$x_{njt+1} = a x_{njt} + b \sum_{\tilde{j} \in J_n \setminus \{j\}} x_{n\tilde{j}t} + c \sum_{\tilde{n} \neq n} \sum_{\tilde{j} \in J_{\tilde{n}}} x_{\tilde{j}\tilde{n}t} + \varepsilon_{njt}$$

with

$$\begin{aligned} a &= (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} - B_1 \frac{1}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{1}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \\ b &= -B_1 \frac{\tau_{within}^{-\theta}}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{\kappa_{within}^{-\theta}}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \\ c &= -B_1 \frac{\tau_{across}^{-\theta}}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{\kappa_{across}^{-\theta}}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \end{aligned}$$

Note that if τ_{within} and κ_{within} are respectively smaller than τ_{across} and κ_{across}, and if the research spillovers are strong enough, then $b > c > 0$.

Let $\bar{x}_{nt}$, $\bar{\varepsilon}_{nt}$ denote the continental average of innovation and disturbances, and $\bar{x}_t$ and $\bar{\varepsilon}_t$ denote the global averages. Further, define $y_{njt} \equiv x_{njt} - \bar{x}_{nt}$ and $u_{njt} \equiv \varepsilon_{njt} - \bar{\varepsilon}_{nt}$ to be a country's deviation from its continental mean and $y_{nt} \equiv \bar{x}_{nt} - \bar{x}_t$ and $u_{nt} \equiv \bar{\varepsilon}_{nt} - \bar{\varepsilon}_t$ to be the continent's deviation from the global mean. The following proposition summarizes the persistence at the country level and the continental level

Proposition 1. *y_{jt} and y_{nt} follow the following processes*

$$y_{njt+1} = (a-b)\, y_{njt} + u_{njt+1}$$

$$y_{nt+1} = [(a-b) + J\,(b-c)]\, y_{nt} + u_{nt+1}$$

In particular, when $b > c > 0$ country-level deviations (y_{nj}) are less persistent than continental deviations (y_n).

Proof. The country law of motion can be expressed as

$$x_{njt+1} = (a-b)\, x_{njt} + bJ\bar{x}_{nt} + c\,(NJ\bar{x}_t - J\bar{x}_{nt}) + \varepsilon_{njt+1}$$

Average over countries in continent n gives

$$\bar{x}_{nt+1} = (a-b)\, \bar{x}_{nt} + bJ\bar{x}_{nt} + c\,(NJ\bar{x}_t - J\bar{x}_{nt}) + \bar{\varepsilon}_{nt+1}$$

Average over continents gives

$$\bar{x}_{t+1} = (a-b)\, \bar{x}_t + bJ\bar{x}_t + c\,(NJ\bar{x}_t - J\bar{x}_t) + \bar{\varepsilon}_{t+1}$$

Subtracting the second from the first gives

$$y_{njt+1} = (a-b)\, y_{njt} + u_{njt+1}$$

Subtracting the third from the second gives

$$\begin{aligned} y_{nt+1} &= (a-b)\, y_{nt} + bJy_{nt} + c\,(-Jy_{nt}) + u_{nt+1} \\ &= [(a-b) + J\,(b-c)]\, y_{nt} + u_{nt+1} \end{aligned}$$

□

The proposition states that when the cost of distance is smaller among countries within the same continent than between countries on different continents, then leadership continental average innovation is more persistent than a county's innovation relative to its continent.

This pattern arises from stronger diffusion within continents than across continents. Note that this type of pattern would not arise from shocks to innovation incentives that are more correlated within continents than between continents, even if such a shock were persistent. Even if such a shock were persistent, it would not generate *more* persistence at the continent level than at the country level.[4]

5 Quantitative Exploration

We present preliminary simulations of the theoretical framework to illustrate its ability to reproduce three key features of the data: (i) innovation is substantially more concentrated than production; (ii) leadership in innovation is more persistent at the continent level than at the country level; (iii) innovation commoves (more) negatively among closer neighbors. We do this by exploring a simple model world economy consisting of 20 countries grouped into two continents with 10 countries each. Geography is parameterized by insight and productivity "discountings", $\kappa_1 = \tau_1 = 2$ and $\kappa_2 = \tau_2 = 5$, for countries on the same and different continents, respectively. Countries are assumed to be symmetric, with the exception of their research productivity. [5]

[4] It is not impossible to get this pattern from correlated shocks to incentives to innovate. It would have to be that the shocks to such incentives had two components, one that was highly correlated within continents but not across continents and persistent, and one that was uncorrelated within continents and short lived.

[5] We use the following common values for the parameter: $\epsilon = 2, \theta = 3, \beta = 0.1, \sigma = 3, B^Y = B^R = B^A = 1, \eta = 0.5, \mu = 1.5$.

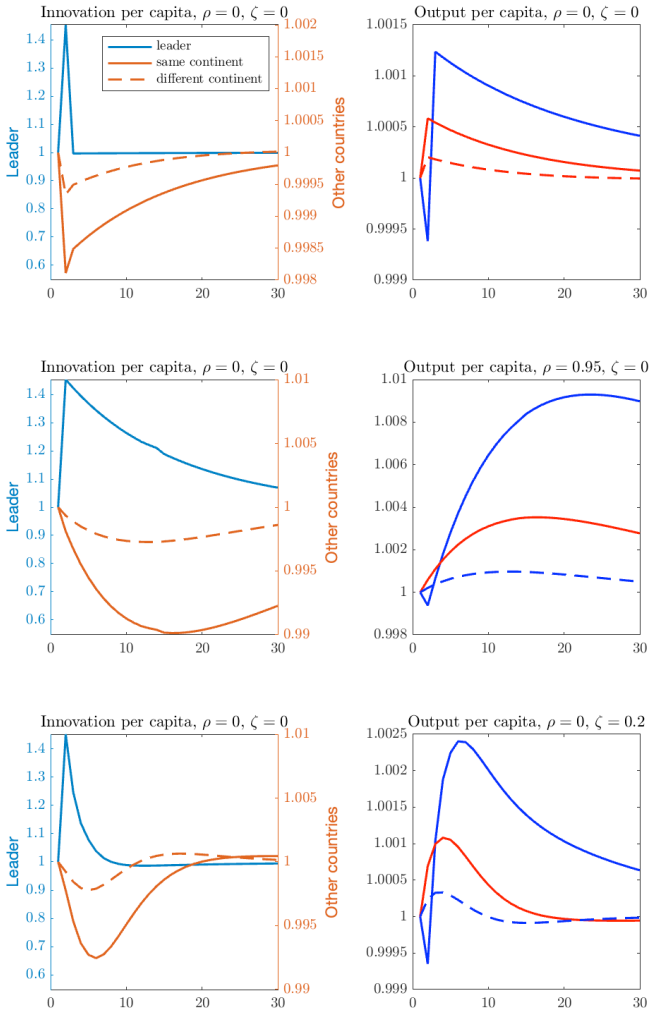
Figure 9 reports the impulse responses of innovation per-capita (left panels) and tangible output per-capita (right panels) to a 10% shock to the innovation productive of a single country. The first row corresponds to the case in which the shock is reverted in one period. The second row assumes that the shock persists: $Z_{1,t} = 0.95 Z_{1,t-1} + 0.05 \cdot 1$, where the steady state productivity is normalized to one, $Z_{\ell,0} = 1$, and $Z_{1,1} = 1.1$. The last row considers the case in which the productivity of the innovation sector is positively affected by past employment in research, $\zeta = 0.2$. The world economy consists of 20 ex-ante symmetric countries, grouped into two continents.

Two key lessons emerge from this exercise. First, innovation in the country receiving the research productivity shock naturally increases, whereas innovation declines for the remaining countries, which reallocate labor to technology development and production. Importantly, innovation declines the most for the closest countries. Secondly, while innovation diverges and becomes highly concentrated (notice the different scales for the variable in the left panels), production is positively affected for all countries.

Next, we consider the evolution of the world economy driven by a history of shocks to the innovation productivity of individual countries. As before, countries are symmetric, with the exception of the realization of the productivity of innovation. Innovation productivity shocks follow an AR(1) in logs: $\log Z_{\ell t} = 0.95 \log Z_{\ell t} + \xi_t$, $\xi_t \sim \mathcal{N}(0, 0.05)$. Figure 10 reports the evolution of innovation and production over 100 periods.

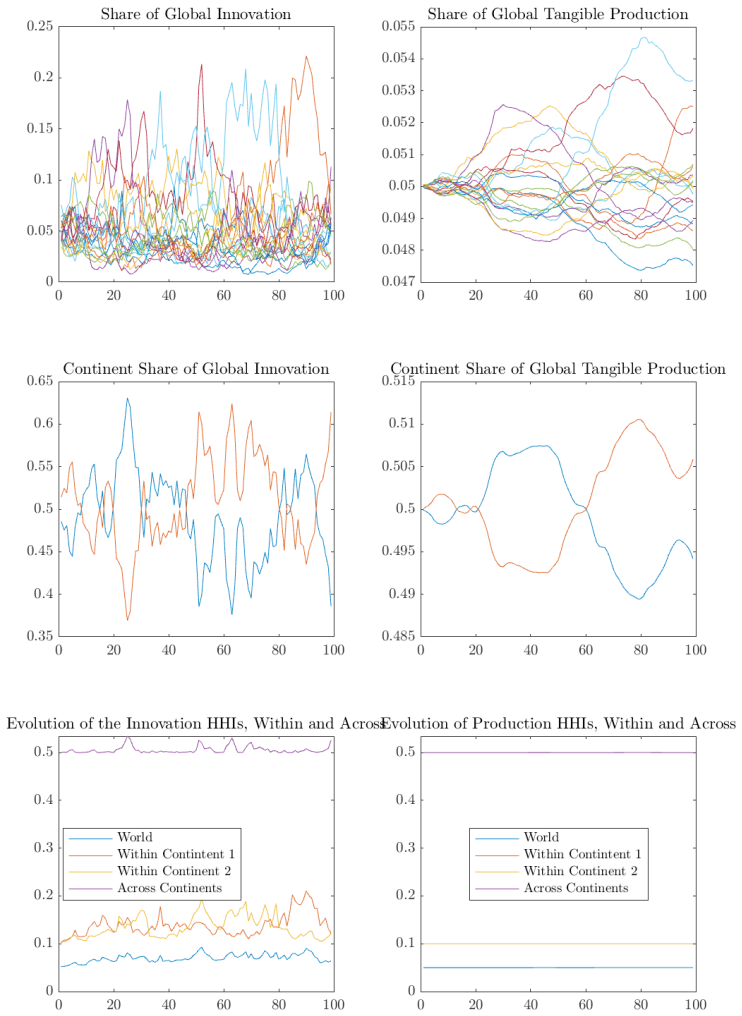
Consistent with the impulse responses in Figure 9 , these simulations further illustrate the ability of the model to generate persistent periods where innovations are highly concentrated in a single leader, while production remains disperse. In addition, the concentration of innovation at the continent level exhibits even more persistence.

Figure 9. Impulse Responses, Two Continents Example



Notes: The figure reports the impulse responses of per-capita innovation (left panels) and tangible output (right panels) to a 10% shock to the innovation productive of a single country. The first row corresponds to the case in which the shock is reverted in one period. The second row assumes that the shock persists: $Z(t) = 0.95Z(t-1) + 0.05 \cdot 1$, where the steady state productivity is normalized to one, $Z(0) = 1$. The last row considers the case in which the productivity of the innovation sector is positively affected by past employment in research, $\zeta = 0.2$. The world economy consists of 20 ex-ante symmetric countries, grouped into two continents.

Figure 10. Evolution of Concentration, Two Continents Example



Notes: The figure reports a simulation of a world economy consist of 20 countries, grouped into two continents. Countries are symmetric with the exception of the realization of the innovation productivity shock. Innovation productivity shocks follow an AR(1) in logs: $\log Z_{\ell t} = 0.95 \log Z_{\ell t} + \xi_t, \xi_t \sim \mathcal{N}(0, 0.05)$.

6 Conclusion

We analyze significant technological innovations over the past millennium to document the patterns of their emergence. The data show that innovation is extraordinarily concentrated geographically, far more so than population or economic output, and that this concentration is broad-based across technological fields and persists even among high-income countries. The centers of innovation shift substantially over the millennium—from Asia and the Middle East in the early centuries, to Europe during the Scientific and Industrial Revolutions, to the United States over the past century. Indeed, innovations per capita can decline over time even among high-income countries. Furthermore, the geographic concentration of innovation does not seem fully explained by agglomeration alone given that the spatial correlation can be negative in the data.

Standard accounts of innovation do not easily explain the full set of facts. We develop a framework combining research, technology diffusion, and a mobile factor to explore their underlying mechanisms. When the mobile factor shifts toward locations with favorable innovation conditions, innovation concentrates while production remains relatively dispersed, since diffusion allows all countries to benefit from innovations made elsewhere. Understanding the determinants of those flows, and how they interact with local fundamentals such as institutions and human capital, seems an important question going forward.

References

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti. 2006. "Distance to frontier, selection, and economic growth." *Journal of the European Economic association*, 4(1): 37–74.

Aghion, Philippe, and Peter Howitt. 1992. "A model of growth through creative destruction." *Econometrica*, 60(2): 323–351.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. 2017. "The rise of American ingenuity: Innovation and inventors of the golden age." Working Paper.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi. 2018. "Dancing with the stars: Innovation through interactions." Working Paper.

Allen, Treb, and Dave Donaldson. 2020. "Persistence and path dependence in the spatial economy." National Bureau of Economic Research.

Alvarez, Fernando E, Francisco J Buera, Robert E Lucas, et al. 2013. "Idea flows, economic growth, and trade." National Bureau of Economic Research.

Asirvatham, Hemanth. 2024. "The dynamo is not the computer." Working Paper.

Asirvatham, Hemanth, Elliott Mokski, and Andrei Shleifer. 2026. "GPT as a measurement tool." Working Paper.

Audretsch, David B, and Maryann P Feldman. 1996. "R&D spillovers and the geography of innovation and production." *American Economic Review*, 86(3): 630–640.

Benhabib, Jess, Jesse Perla, and Christopher Tonetti. 2014. "Catch-up and fall-back through innovation and imitation." *Journal of Economic Growth*, 19(1): 1–35.

Bergeaud, Antonin, Ruveyda Nur Gozen, and John Van Reenen. 2026. "Mapping technological trajectories: Evidence from two centuries of patent data." Working Paper.

Bernard, Andrew B, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum. 2003. "Plants and productivity in international trade." *American Economic Review*, 93(4): 1268–1290.

Brown, James R, Steven M Fazzari, and Bruce C Petersen. 2009. "Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom." *Journal of Finance*, 64(1): 151–185.

Buera, Francisco J, and Ezra Oberfield. 2020. "The global diffusion of ideas." *Econometrica*, 88(1): 83–114.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2022. "Knowledge diffusion, trade, and innovation across countries and sectors." *American Economic Journal: Macroeconomics*, 14(1): 104–145.

Carlino, Gerald, and William R Kerr. 2015. "Agglomeration and innovation." In *Handbook of Regional and Urban Economics*. Vol. 5, 349–404. Elsevier.

Cavalcanti, Tiago VV, Chryssi Giannitsarou, and Charles R Johnson. 2017. "Network cohesion." *Economic Theory*, 64(1): 1–21.

Chikis, Craig A, Benny Kleinman, and Marta Prato. 2025. "The geography of innovative firms." Working Paper.

Comin, Diego, and Bart Hobijn. 2004. "Cross-country technology adoption: Making the theories face the facts." *Journal of Monetary Economics*, 51(1): 39–83.

Comin, Diego, and Bart Hobijn. 2010. "An exploration of technology diffusion." *American Economic Review*, 100(5): 2031–2059.

Comin, Diego, and Martí Mestieri. 2018. "If technology has arrived everywhere, why has income diverged?" *American Economic Journal: Macroeconomics*, 10(3): 137–178.

Comin, Diego, and Ramana Nanda. 2019. "Financial development and technology diffusion." *IMF Economic Review*, 67(2): 395–419.

Comin, Diego, William Easterly, and Erick Gong. 2010. "Was the wealth of nations determined in 1000 BC?" *American Economic Journal: Macroeconomics*, 2(3): 65–97.

Eaton, Jonathan, and Samuel Kortum. 1999. "International technology diffusion: Theory and measurement." *International Economic Review*, 40(3): 537–570.

Giroud, Xavier, Ernest Liu, and Holger M Mueller. 2026. "Innovation spillovers across U.S. tech clusters." *Journal of Financial Economics*. Forthcoming.

Golub, Benjamin, and Matthew O Jackson. 2012. "How homophily affects the speed of learning and best-response dynamics." *The Quarterly Journal of Economics*, 127(3): 1287–1338.

Grossman, Gene M, and Elhanan Helpman. 1991. "Quality ladders and product cycles." *The Quarterly Journal of Economics*, 106(2): 557–586.

Hsu, Po-Hsuan, Xuan Tian, and Yan Xu. 2014. "Financial development and innovation: Cross-country evidence." *Journal of Financial Economics*, 112(1): 116–135.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic localization of knowledge spillovers as evidenced by patent citations." *Quarterly Journal of Economics*, 108(3): 577–598.

Jones, Charles I. 1995. "R&D-based models of economic growth." *Journal of Political Economy*, 103(4): 759–784.

Jovanovic, Boyan, and Rafael Rob. 1989. "The growth and diffusion of knowledge." *The Review of Economic Studies*, 56(4): 569–582.

Kalyani, Aakash, Nicholas Bloom, Marcela Carvalho, Tarek Hassan, Josh Lerner, and Ahmed Tahoun. 2025. "The diffusion of new technologies." *Quarterly Journal of Economics*, 140(2): 1299–1365.

Kerr, William R, and Ramana Nanda. 2015. "Financing innovation." *Annual Review of Financial Economics*, 7: 445–462.

King, Robert G, and Ross Levine. 1993. "Finance and growth: Schumpeter might be right." *Quarterly Journal of Economics*, 108(3): 717–737.

Kortum, Samuel S. 1997. "Research, patenting, and technological change." *Econometrica: Journal of the Econometric Society*, 1389–1419.

Lind, Nelson, and Natalia Ramondo. 2023. "Global innovation and knowledge diffusion." *American Economic Review: Insights*, 5(4): 494–510.

Liu, Ernest, and Song Ma. 2021. "Innovation networks and R&D allocation." Working Paper.

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

Romer, Paul M. 1990. "Endogenous technological change." *Journal of Political Economy*, 98(5, Part 2): S71–S102.

Trouvain, Florian. 2024. "Technology adoption, innovation, and inequality in a global world."

Internet Appendix

IA1 Processing Wikipedia Titles for Technological Innovations

This appendix describes in detail how we obtain and process Wikipedia titles on technological inventions used in Section 3.

IA1.1 Wikipedia Titles on Technological Inventions

We extract English Wikipedia titles from the Wikimedia English Wikipedia Dumps and filter them in two steps to identify titles related to technological inventions.

Step 1 We first do batch filtering to obtain a broad set of technology-related titles using GPT-4o-mini. The LLM is instructed to identify Wikipedia titles corresponding to technological inventions. Wikipedia titles are processed in batches of 50, and the LLM returns only those classified as technological inventions.

```
Instructions:
- You are tasked with filtering a list of Wikipedia article titles to
include only those of technological inventions. A technological invention is
a novel device, method, process, or system that applies scientific or
engineering principles to solve a problem or improve efficiency. These
inventions typically introduce new functionalities, enhance existing
technologies, or create entirely new categories of tools or systems.
- Focus ONLY on technological inventions, which are devices, systems, and
components with novel technological advancements, such as ''wheels'',
''artificial intelligence'', ''eletric vehicle'', ''mobile phone'', ''Watt
steam engine'', etc.
- EXCLUDE names of people, movies, books, songs, food, animals, places,
companies, organizations, social movements, slogans, cultural phrases, and
any non-technology items.
- EXCLUDE specific models, product names, versions of weapons and vehicles,
unless they represent a significant technological advancement (e.g.,
''iPhone'' is a technological invention, but ''iPhone 13'' is not. Airplane
is a technological invention, but ''Boeing 747'' is not).
- Do NOT include scientific phenomena unless they describe a concrete
technological invention or device.
```

```
- If multiple titles are extremely similar (such as variations in
capitalization, punctuation, or wording), ONLY include ONE representative
title.

For each title, first assess if it represents a technological invention.
Then, return ONLY the valid titles (keeping the original format of input,
including the ''_'' character).

Respond in a JSON array of strings like this:
[''Title 1'', ''Title 2'', ''Title 3'']

Titles:
{titles_list}
```

Step 2 We validate each Wikipedia title from Step 1 using its Wikipedia summary to obtain the final set of technological inventions. For each title, we use the Wikipedia API library to retrieve the corresponding English Wikipedia summary. We then supply both the title and the summary to Deepseek-Reasoner model, which is prompted to determine whether the article indeed describes a technological invention, following the same definition used in Step 1. The LLM's responses indicate whether the Wikipedia title should be retained.

```
Instructions:
- You are tasked with determining whether a given Wikipedia article is about
a technological invention based on the title and summary provided. A
technological invention is a novel device, method, process, or system that
applies scientific or engineering principles to solve a problem or improve
efficiency. These inventions typically introduce new functionalities,
enhance existing technologies, or create entirely new categories of tools or
systems.
- EXCLUDE specific models, product names, versions of weapons and vehicles,
unless they represent a significant technological advancement (e.g.,
''iPhone'' is a technological invention, but ''iPhone 13'' is not. Airplane
is a technological invention, but ''Boeing 747'' is not). Example of titles
that should be EXCLUDED: ''AJS_Model_16'', ''AMD_2900'',
''A340_(aircraft)'', ''A.T_Mine_E.P._Mark_II''.

Wikipedia Article Title: {title}
```

```
Wikipedia Article Summary:
\''\''\''{summary}\''\''\''

Based on the above summary, determine if this article is about a
technological invention.
Respond ONLY with a JSON object in the following format:
{{''is_valid'': true}}
or
{{''is_valid'': false}}
```

IA1.2 Technology Information

We use the following prompt to summarize basic information about the technological inventions with Gemini-2.0-Flash, including the time and location of the invention, as well as the type of inventor (i.e., private company, individual inventor, government, or university/non-profit).

```
Instructions:
You are an expert in the history of technologies. Your task is to answer
questions about when, where, who invented the technology, and which category
of application it best fits into, based on the given information about the
technology and your domain knowledge.

Provide the following information along with a brief reason (less than 20
words) for each:

1) When the technology first became workable. If the exact year is
unavailable or uncertain, provide the closest approximation. Be as accurate
as possible. Include ''BCE'' in your response if the date is before 0
CE/very ancient.

2) Where it was first created (country where the first workable version was
created).

3) Who invented it (choose one of these categories by the type of inventor):
- ''Private Company'': For-profit company/companies or people working for
these companies developed the technology (e.g. Apple, Google);
```

- ''University/Non-Profit'': Non-profit institution or organization developed it, such as universities, non-profit institutions, open-source projects (e.g. UC Berkeley, SRI, Mozilla);
- ''Government Project'': Developed via government-funded or military project (e.g. NASA, the Apollo program, the Manhattan Project);
- ''Individual Inventor'': Developed by an individual or a small team of inventors unaffiliated with large institutions or companies (e.g. Nikola Tesla, Philo Farnsworth);

Note: Some inventions involve collaboration. When multiple parties are involved, choose the category that best reflects the majority of developers or the most responsible entity. Aim for the most accurate classification based on available evidence. When inventors are not explicitly mentioned, use your best judgment to classify the technology based on the context and purpose of the invention
For example:
- A collaboration between private companies is still classified as ''Private Company''.
- A small team of unaffiliated inventors is still ''Individual Inventor''.
- If an individual works under a government-funded initiative, classify it as ''Government Project''.
- If it was published by an industry association, it should be either ''Private Company'' or ''University/Non-Profit'' depending on the inventor and purpose of the technology. For example, SAE (Society of Automotive Engineers) is mostly ran by engineers from private companies, so it should be classified as ''Private Company''. W3C (World Wide Web Consortium) is ran by university researchers, so it should be classified as ''University/Non-Profit''.
- If it was designed for commercial use or production efficiency, it's likely ''Private Company''.
- If it served military, national defense, or public interest, it's likely ''Government Project''.
- If it was developed for academic research or knowledge advancement, it's likely ''University/Non-Profit''.
- If it was created by individuals not working for large institutions, it's likely ''Individual Inventor''.

```
4) The 'domain_of_application' (choose exactly one from the list below):
   [''Agriculture, Forestry, Fishing and Hunting'', ''Mining, Quarrying, and
   Oil and Gas Extraction'', ''Utilities'', ''Construction'',
   ''Manufacturing'', ''Wholesale Trade'', ''Retail Trade'',
   ''Transportation and Warehousing'', ''Information'', ''Finance and
   Insurance'', ''Real Estate and Rental and Leasing'', ''Professional,
   Scientific, and Technical Services'', ''Health Care and Social
   Assistance'', ''Military and Defense'', ''Basic Research'']

Importantly: Use only the listed categories above. Do not invent new
categories.

Technology Name: {title}

Wikipedia Page:
\''\''\''{wiki_text[:5000]}\''\''\''

Return ONLY a JSON object in the following format:
{{
  ''technology'': ''Title'',
  ''when'': year or ''Not Available'',
  ''when_reason'': ''Brief reason for chosen year'',
  ''where'': country name or ''Not Available'',
  ''where_reason'': ''Brief reason for chosen country'',
  ''who'': one of the 4 listed categories,
  ''who_reason'': ''Brief reason for chosen category'',
  ''domain_of_application'': one of the 15 listed categories,
  ''domain_reason'': ''Brief reason for chosen domain''
}}

Example Output #1:
{{
  ''technology'': ''iPhone'',
  ''when'': ''2007'',
  ''when_reason'': ''Apple released the first iPhone in 2007.'',
  ''where'': ''USA'',
  ''where_reason'': ''Apple developed the iPhone in the United States.'',
```

```
  ''who'': ''Private Company'',
  ''who_reason'': ''Developed by Apple, which is a for profit company.'',
  ''domain_of_application'': ''Information'',
  ''domain_reason'': ''Primarily used for communication and computing.''
}}

Example Output #2:
{{
  ''technology'': ''CRISPR'',
  ''when'': ''2012'',
  ''when_reason'': ''First practical demonstration of gene editing was in
  2012.'',
  ''where'': ''USA'',
  ''where_reason'': ''Initial research conducted at UC Berkeley.'',
  ''who'': ''University Lab'',
  ''who_reason'': ''Developed at UC Berkeley by academic researchers.'',
  ''domain_of_application'': ''Health Care and Social Assistance'',
  ''domain_reason'': ''Primarily applied to medicine and biotechnology.''
}}

Use this JSON schema:
{{
  ''technology'': ''str'',
  ''when'': ''str'',
  ''when_reason'': ''str'',
  ''where'': ''str'',
  ''where_reason'': ''str'',
  ''who'': ''str'',
  ''who_reason'': ''str'',
  ''domain_of_application'': ''str'',
  ''domain_reason'': ''str''
}}
```

IA1.3 State Information of U.S. Technological Innovations

Our main analyses and the LLM query in Section IA1.2 collect the location of the innovation at the country level. For the U.S. state-level analyses, we need additional information about the state of invention for U.S. innovations. We use Gemini-2.0-Flash with grounded web search enabled to supplement information in Wikipedia.

```
You are an expert in the history of technologies.

Task:
Here is a technology invented in the U.S. Please determine the state where
it was first invented.

Rules:
- Use the Wikipedia page of the technology title below as the primary
source.
- You may use grounded web search if needed.
- The answer must be a two-letter U.S. state code (e.g., CA, NY, MA).
- If the state cannot be determined with reasonable confidence, write "Not
Available".
- Do NOT guess.

Technology: {technology}
Year: {year}

Wikipedia Content:
\"\"\"{wiki_text[:5000]}\"\"\"

Return STRICT JSON only:
{{
  "technology": "{technology}",
  "year": "{year}",
  "state": "str or Not Available",
  "source": "Wikipedia or Wikipedia + web"
}}
```

IA2 Processing Wikipedia Titles for Artwork

This appendix describes in detail how we obtain and process Wikipedia titles on artwork used in Section 3.

IA2.1 Wikipedia Titles on Artwork

We extract English Wikipedia titles from the Wikimedia English Wikipedia Dumps and filter them in two steps to identify titles related to artwork.

Step 1 Unlike technological inventions, which span a wide and heterogeneous range of topics that are difficult to delineate with simple rules, artwork on Wikipedia is systematically organized into well-defined categories corresponding to artistic genres and media (e.g., "paintings," "novels," "films," "sculptures"). We exploit this structure by performing keyword-based filtering on the Wikipedia categories associated with each title to obtain a broad set of artwork-related titles. We retain titles whose categories contain at least one of the following keywords: painting, sculpture, photography, film, music, literature, performing arts, architecture, installation, digital art, ceramics, opera, ballet, theatre, fresco, mural, portrait, landscape painting, still life, oil painting, watercolor, ink painting, miniature painting, illumination, icon painting, altarpiece, triptych, diptych, statue, bust, relief, carving, bronze sculpture, marble sculpture, wood sculpture, terracotta, monument, photograph, photographic print, daguerreotype, photomontage, decorative arts, visual arts, fine arts, craft, folk art, applied arts, design, iconography, artifact, masterpiece, among others.

Step 2 We validate each Wikipedia title from Step 1 using its Wikipedia categories and summary to obtain the final set of artwork. For each title, we use the Wikipedia API library to retrieve the corresponding English Wikipedia summary. We then supply the title, categories, and summary to the Deepseek-Chat model, which is prompted to determine whether the article indeed describes a single, concrete artwork. The LLM's responses indicate whether the Wikipedia title should be retained.

```
You are a strict JSON-only classifier.

Task:
Determine whether the given Wikipedia entry refers to a single, concrete
artwork.

Definition:
```

```
A specific creative work in painting, sculpture, photography, film, music,
literature, performing arts, architecture, installation, digital art,
ceramics, opera, ballet, theatre, or other artistic media.
Examples: ''Mona Lisa'', ''The Starry Night'', ''Hamlet'', ''The Rite of
Spring'', ''Forrest Gump'', ''Bohemian Rhapsody'', ''Swan Lake''.

Exclude:
- People, concepts, events, movements, genres, or general terms (e.g.
Romanticism, Jazz, Painting in China).
- Organizations, companies, locations, periods.
- Generic categories that are not a specific, notable artwork.

Return ONLY a JSON object like:
{{''is_valid'': true, ''explanation'': ''short reason''}}

Title: {title}

Categories:
\''\''\''{categories}\''\''\''

Summary:
\''\''\''{summary}\''\''\''
```

IA2.2 Artwork Information

We use the following prompt to summarize basic information about the artwork with Gemini-2.0-Flash, including the genre, creator, year and location of creation.

```
You are an expert cataloguer of artworks. Your task is to extract structured
facts about the artwork's title, artist, creation date and location, and
genre, based on the given information about the artwork and your domain
knowledge.

Provide the following information:

1) Title of the artistic piece
```

```
2) Genre of the artwork: Choose EXACTLY ONE genre from this fixed set (do
not invent new labels):
[''Painting'', ''Sculpture'', ''Photography'', ''Film'', ''Music'',
''Literature'', ''Performing Arts'', ''Architecture'', ''Installation'',
''Digital Art'', ''Ceramics'', ''Opera'', ''Ballet'', ''Theatre'',
''Other'']
If the work spans multiple forms, choose the best primary genre.

3) Creator(s) of the artwork: Include the primary credited creator(s). If
multiple, join names with ''; ''

4) Creator's country of origin: If multiple locations are involved, choose
the most significant one. If unclear, return ''Not Available''. If multiple
creators, choose the country of origin of the most primary creator. If equal
prominence, choose the country associated with where that creator conducted
most of their career-defining work. If still unclear, return ''Mixed''.

5) Year of creation: If the exact year is unavailable or uncertain, provide
the closest approximation. Be as accurate as possible. Include ''BCE'' in
your response if the date is before 0 CE/very ancient.

6) Country of creation: The country where the work was principally
created/produced. If multiple locations, pick the primary
production/creation location. If unclear, ''Not Available''.

Title: {title}

Wikipedia Page:
\''\''\''{wiki_text[:5000]}\''\''\''

Output strictly as JSON (no backticks, no preamble, no trailing text),
schema:
{{
  ''piece_name'': ''str'',
  ''genre'': ''str'',
  ''creator'': ''str'',
  ''creator_country_of_origin'': ''str'',
```

```
  ''year_of_creation'': ''str'',
  ''country_of_creation'': ''str''
}}
```

IA2.3 State Information of U.S. Literature

We also supplement the main query in Section IA2.2 with the state most closely associated with each work's creation for U.S. literature. We use Gemini-2.0-Flash with grounded web search enabled to supplement Wikipedia information, and apply the following priority order: (1) the state where the author wrote or completed the work, (2) the state where the work was first published, and (3) the state where the author was based at the time of writing.

```
You are an expert in American literary history.

Task:
Here is a literary work created in the U.S. Please determine the U.S. state
most closely associated with its creation - specifically where the author
wrote it, where it was first published, or where the author was based at the
time of writing.

Rules:
- Use the Wikipedia summary of the literary work provided below as the
primary source.
- You MUST perform a grounded web search to supplement or verify the
information, especially when the summary does not clearly indicate the
state.
- The answer must be a two-letter U.S. state code (e.g., CA, NY, MA).
- Base the state on the following priority order:
    1. The state where the author wrote or completed the work.
    2. The state where the work was first published (location of the
    publisher).
    3. The state where the author was living or based at the time of
    writing.
- If multiple states are equally plausible, choose the one most directly
linked to the creation of the work.
- If the state cannot be determined with reasonable confidence, write "Not
Available".
```

```
- Do NOT guess.

Literary Work: {title}
Year: {year}

Wikipedia Summary:
\"\"\"{str(summary)[:5000]}\"\"\"

Return STRICT JSON only:
{{
  "wiki_title": "{title}",
  "year": "{year}",
  "state": "str or Not Available",
  "source": "Wikipedia or Wikipedia + web"
}}
```

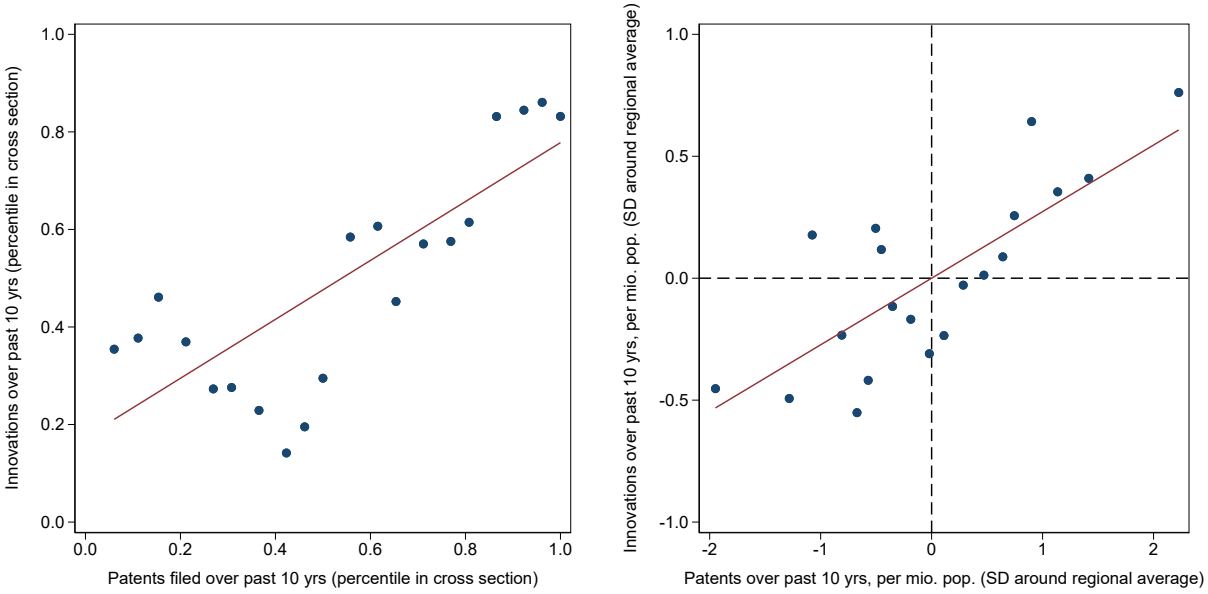
IA3 Additional Results

IA3.1 Additional Figures and Tables

Figure IA1. Innovations Based on Wikipedia Titles and Patents

(a) Relation Across Regions

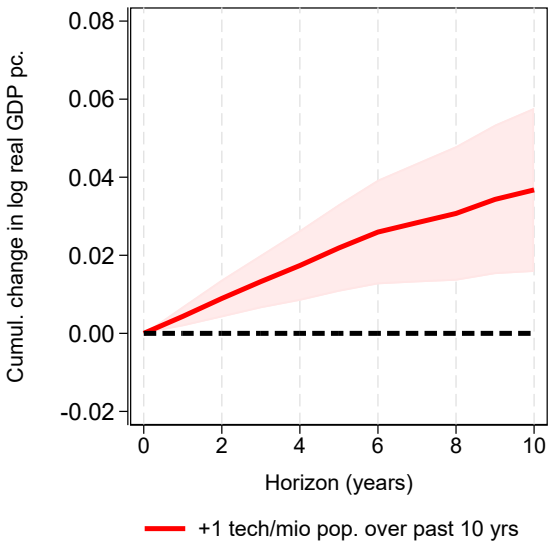
(b) Relation Within Region



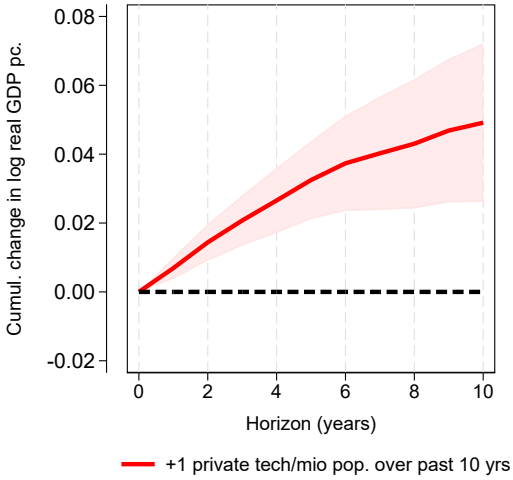
Notes: The figure shows binned scatter plots of regional patents and innovations over the past decade between 1970 and 2000. The left panel plots the data transformed into annual cross sectional percentiles. The right panel plots the temporal deviations from regional averages normalized by regional standard deviations. The patent data comes from PATSTAT. The underlying 28 regions are made of four African, four American, eight Asian, ten European regions and Australia.

Figure IA2. Relationship Between Innovations and Growth

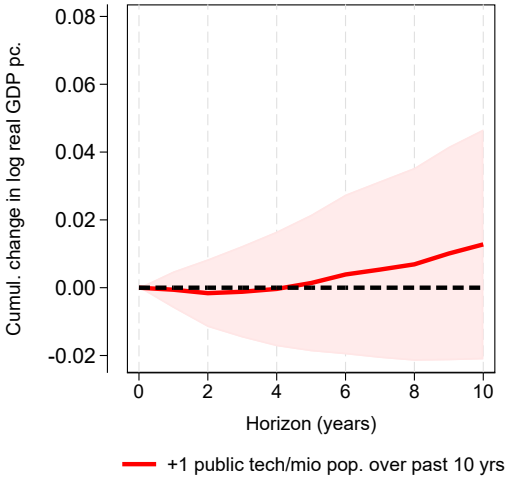
(a) Total Innovations



(b) Private Innovations



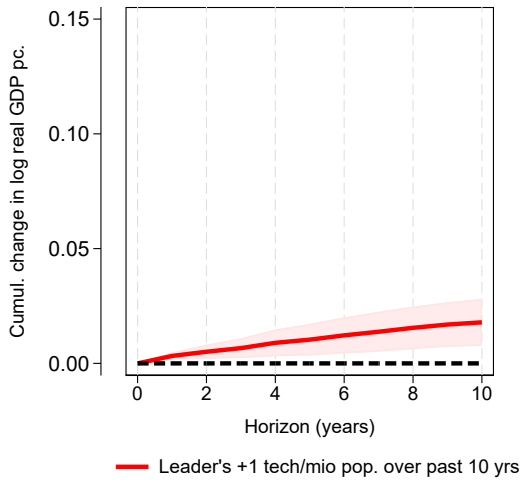
(c) Public Innovations



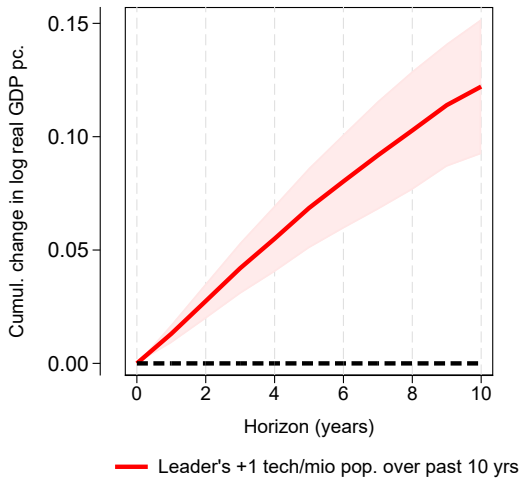
Notes: The figure visualizes changes in a country's future real GDP per capita predicted from increases in innovation activity estimated from the following local projections: $\Delta_h \log(\text{real GDP pc}_{i,t+h}) = \gamma_i + \tau_t + \beta z_{i,t} + \boldsymbol{\delta x}_{i,t} + \epsilon_{i,t}$ where i indexes the country, t the year and h the prediction horizon. The control vector $x_{i,t}$ contains ten lags and the contemporaneous annual real GDP per capita growth as well as the level of log real GDP per capita to control for catch up growth. The sample covers years 1500 to 2000. Data on real GDP per capita comes from Maddison, interpolating historical observations. In Panel (a), $z_{i,t}$ is the number of total innovations conceived over the past ten years in country i by year t relative to population in year t. In Panel (b), $z_{i,t}$ is the number of private innovations conceived over the past ten years in country i by year t relative to population in year t. In Panel (c), $z_{i,t}$ is the number of public innovations conceived over the past ten years in country i by year t relative to population in year t. Shaded areas mark 95% confidence intervals based on standard errors clustered twoway at the country and year level.

Figure IA3. Relationship Between Leader's Innovations and Local Growth

(a) Sample 1500-1899

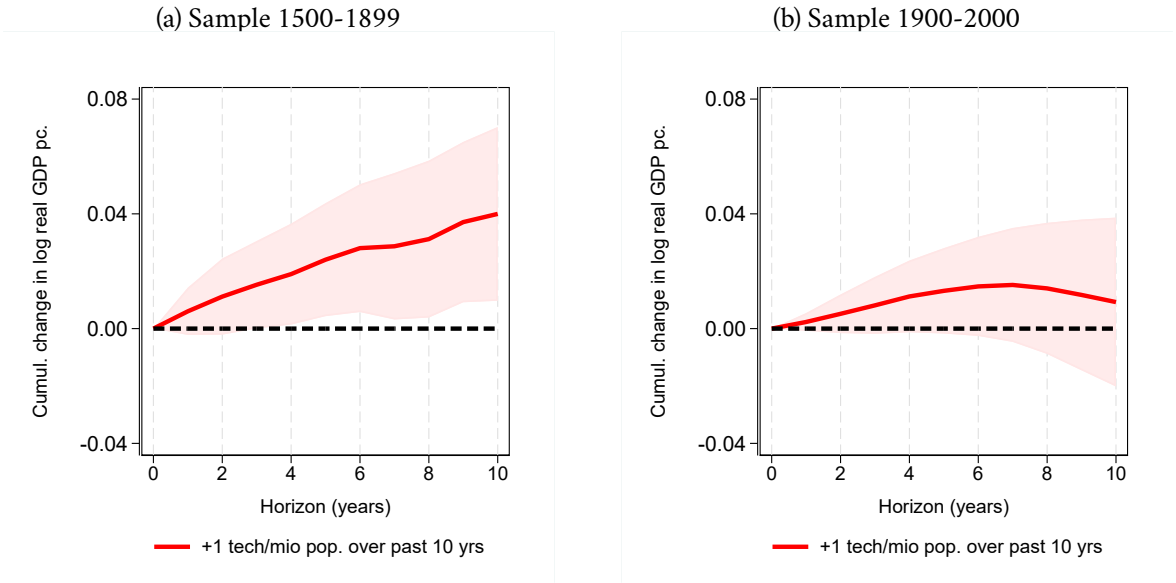


(b) Sample 1900-2000



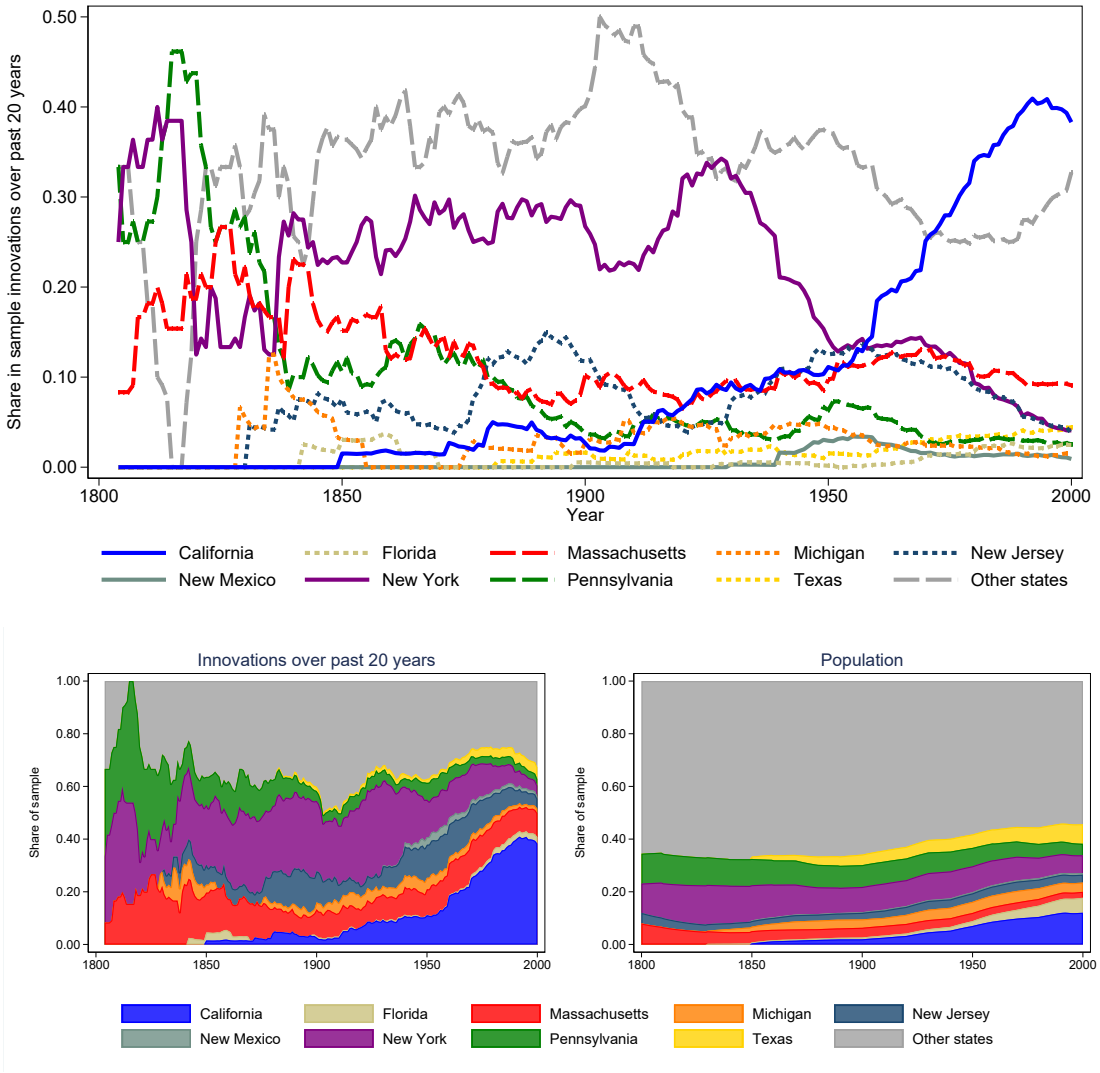
Notes: The figure visualizes changes in a country's future real GDP per capita predicted from increases in innovation activity of the global leader. Estimates come from the following local projections: $\Delta_h \log(\text{real GDP pc}_{i,t+h}) = \gamma_i + \beta z_t + \boldsymbol{\delta x}_{i,t} + \epsilon_{i,t}$ where i indexes the country, t the year and h the prediction horizon. The control vector $x_{i,t}$ contains the number of local innovations over the past ten years relative to local population, ten lags and the contemporaneous annual real GDP per capita growth as well as the level of log real GDP per capita to control for catch up growth. z_t is the number of total innovations conceived over the past ten years relative to population in year t by the global innovation leader. The global innovation leader in year t is defined as having the most innovations over the past ten years relative to population. Panel (a) shows estimates obtained for the sample 1500-1899. Panel (b) shows estimates obtained for the sample 1900-2000. Data on real GDP per capita comes from Maddison, interpolating historical observations. Shaded areas mark 95% confidence intervals based on standard errors clustered twoway at the country and year level.

Figure IA4. Relationship Between Innovations and Growth Over Time



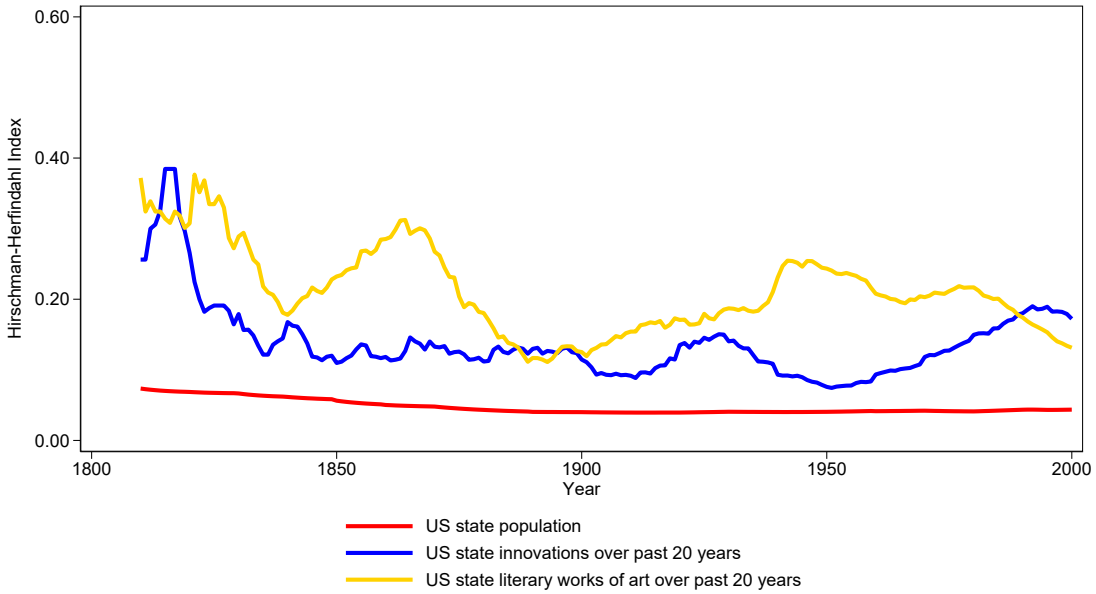
Notes: The figure visualizes changes in a country's future real GDP per capita predicted from increases in innovation activity. Estimates come from the following local projections: $\Delta_h \log(\text{real GDP pc}_{i,t+h}) = \gamma_i + \beta z_{i,t} + \boldsymbol{\delta x}_{i,t} + \epsilon_{i,t}$ where i indexes the country, t the year and h the prediction horizon. The control vector $x_{i,t}$ contains the number of local innovations over the past ten years relative to local population, ten lags and the contemporaneous annual real GDP per capita growth as well as the level of log real GDP per capita to control for catch up growth. $z_{i,t}$ is the number of total innovations conceived over the past ten years relative to population in country i by year t. Panel (a) shows estimates obtained for the sample 1500-1899. Panel (b) shows estimates obtained for the sample 1900-2000. Data on real GDP per capita comes from Maddison, interpolating historical observations. Shaded areas mark 95% confidence intervals based on standard errors clustered twoway at the country and year level.

Figure IA5. Innovation Concentration at the U.S. State Level



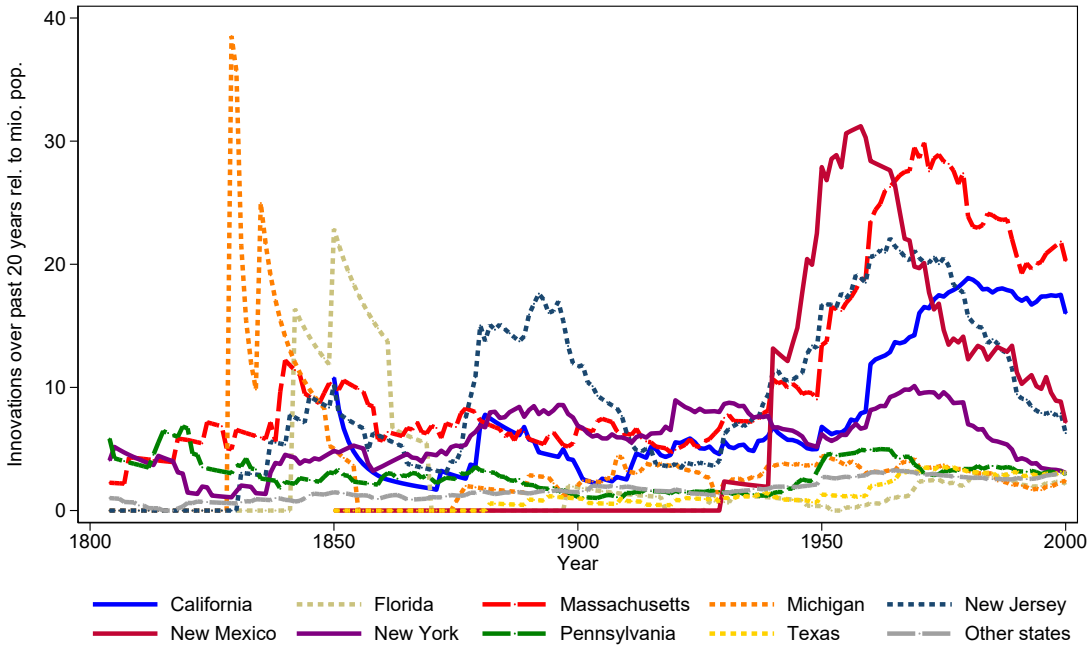
Notes: The figure illustrates the weight of U.S. states in national innovation and population between years 1800 and 2000. The top panel shows the share of states in national innovations. The bottom panel shows stacked shares for national innovation and population shares. Innovations are measured as cumulative sum over the past 20 years, to address lumpiness in the time series in early parts of the sample.

Figure IA6. Hirschman-Herfindahl Indices at the U.S. State Level



Notes: The figure plots Hirschman-Herfindahl concentration indices of innovation and population at the U.S. state level between years 1800 and 2000. Innovations are measured as cumulative sum over the past 20 years, to address lumpiness in the time series in early parts of the sample.

Figure IA7. Innovations per Capita by U.S. State



Notes: The figure plots the number of innovations conceived in a given U.S. state relative to that state's population between years 1800 and 2000. Innovations are measured as cumulative sum over the past 20 years, to address lumpiness in the time series in early parts of the sample.

Table IA1 – Correlates of Total Innovation per Population

	(1)	(2)	(3)	(4)
Log population	0.233	0.455	0.411	1.638
	(0.409)	(0.905)	(0.380)	(1.332)
Log real GDP p.c.	0.289	0.511	1.474***	0.381
	(0.257)	(0.299)	(0.440)	(0.551)
Universities p.c.	0.743**	-0.347	0.576**	-0.481
	(0.270)	(0.363)	(0.245)	(0.297)
Loans rel. to GDP	0.345	0.005	0.426	-0.114
	(0.295)	(0.214)	(0.270)	(0.170)
Stock market cap. rel. to GDP	0.358**	0.213*	0.322	0.148
	(0.156)	(0.116)	(0.217)	(0.122)
Constant	0.272	0.955	-0.844	-0.166
	(0.497)	(0.873)	(0.641)	(1.771)
Country FE		✓		✓
Year FE			✓	✓
R^2	0.12	0.21	0.28	0.36
within R^2	0.12	0.02	0.15	0.01
Obs	1650	1650	1650	1650

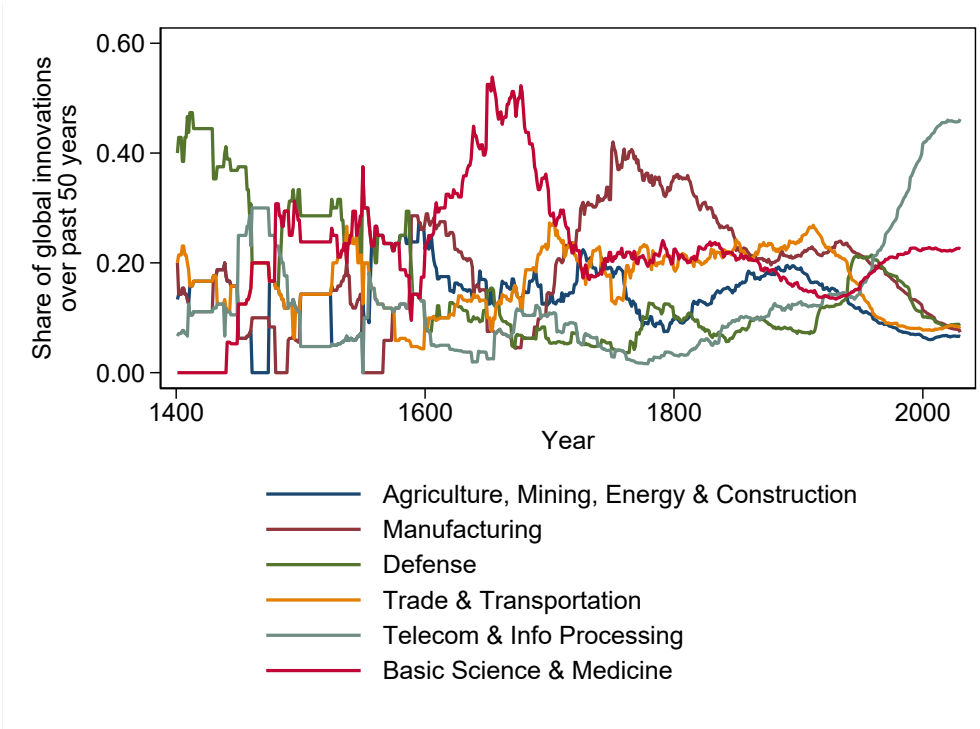
Notes: This table presents estimates from OLS regressions of innovations per capita on social, economic and financial development indicators. Dependent variable and regressors are standardized such that coefficients' magnitude can be compared to gauge relative explanatory power. Specifications shown in different columns differ only in their fixed effects setup. Given limited data availability on financial development indicators, the sample comprises 17 advanced economies between years 1870 and 2000. The 17 countries are: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and the United States. Standard errors are double clustered on levels of countries and years and shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

IA3.2 Decomposing Local and Global Factors in Innovation Activity

Main domain of global innovation activity changes over time. Figure IA8 illustrates the evolution from defense-focused innovation activity before 1600, followed by the scientific revolution, the industrial revolution starting in the 18th century, the automotive revolution in the early 20th century and finally the spurt in innovation in telecommunication and information processing.

Driven by global domain-specific component or the fact that globally leading countries specialized in different domains at different times?

Figure IA8. Decomposition of Global Innovation by Domain



Estimate following fixed effects regression of the logarithm of total innovations i in domain d in country c at year t:

$$\log(i_{c,t,d}) = \underbrace{\gamma_{d,t}}_{\text{global domain component}} + \underbrace{\gamma_{c,t}}_{\text{local component}} + e_{c,t,d} \tag{IA1}$$

$R^2 = 0.88$. Model implies that growth of innovation activity can be decomposed into a global domain component, a local country-specific component and an idiosyncratic domain-country component, i.e.,

$$\Delta \log(i_{c,t,d}) = \Delta\gamma_{t,d} + \Delta\gamma_{c,t} + \Delta e_{c,t,d} \tag{IA2}$$

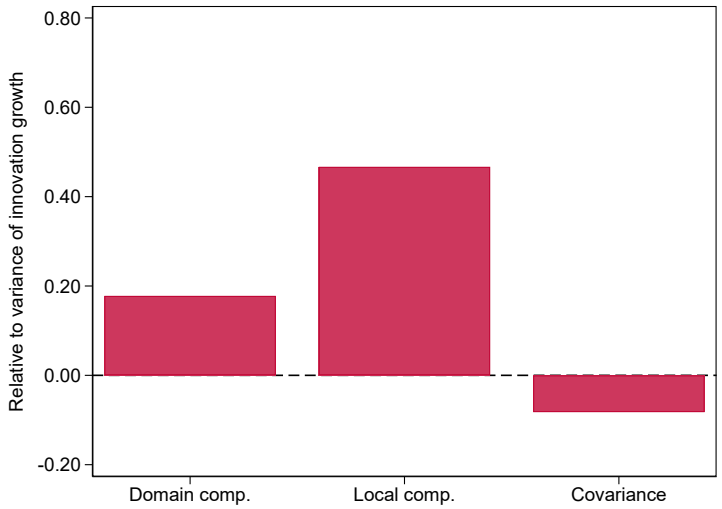
Variance decomposition:

$$Var\Big(\Delta \log(i_{c,t,d})\Big) = Var(\Delta\gamma_{t,d}) + Var(\Delta\gamma_{c,t}) + 2Cov(\Delta\gamma_{t,d}, \Delta\gamma_{c,t}) + Var(\Delta e_{c,t,d}) \tag{IA3}$$

Visualized in Figure IA9. Pattern very similar when done for each domain separately.

Moreover, can construct *counterfactual* innovation series muting the global (country) component of innovation. Figure IA10 shows time series for the global component for each domain. Figure IA11a shows counterfactual innovation activity in major countries since 1600 if the global domain component was muted. Figure IA11b shows counterfactual innovation activity in major countries since 1600 if the local country component was muted. Overall, global domain-specific innovation shocks matter

Figure IA9. Variance Decomposition



quantitatively only for small countries.

Figure IA10. Global Components

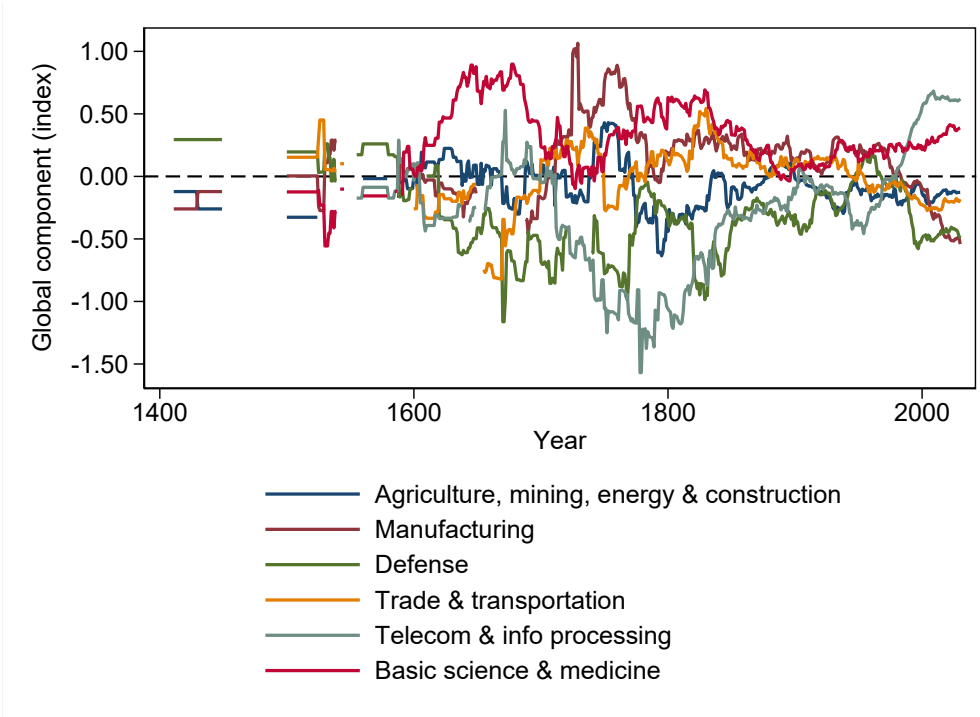
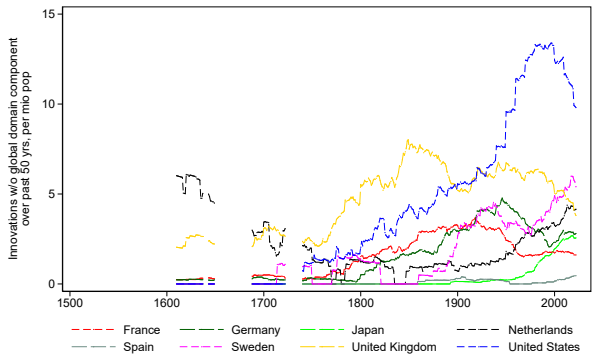
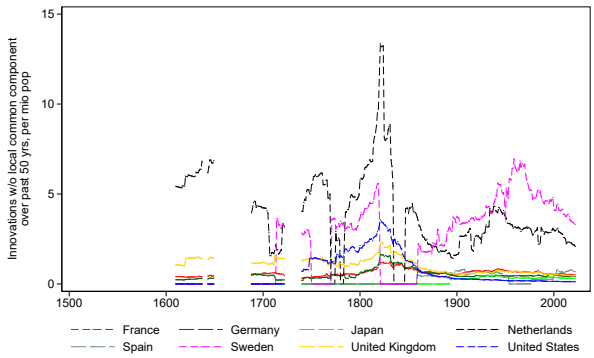


Figure IA11. Global vs Local Components for Aggregate Innovation

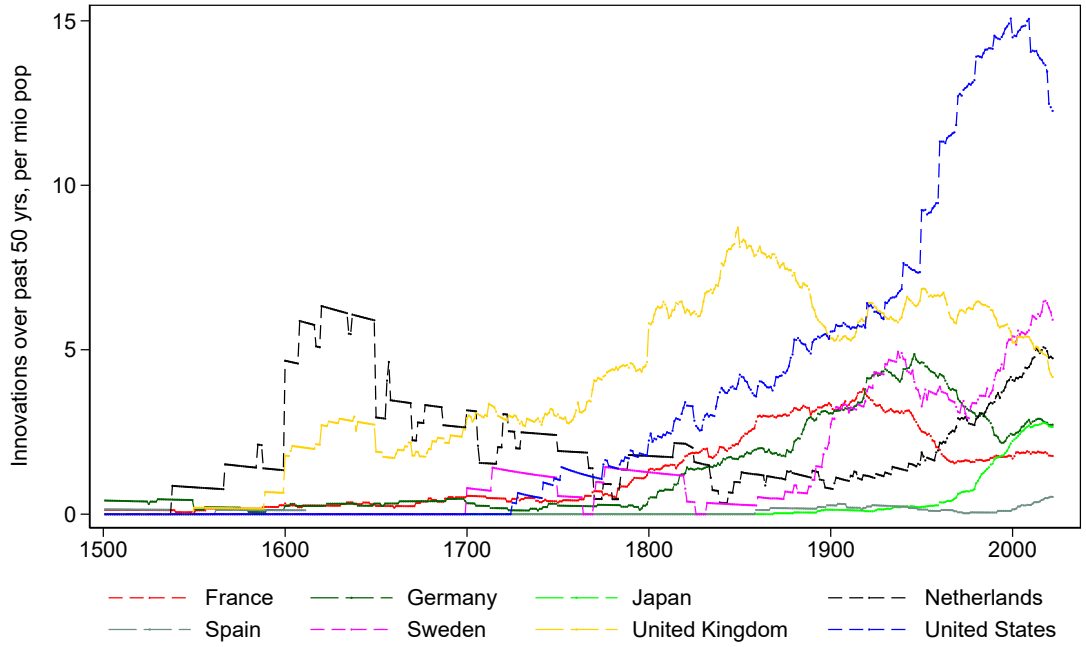
(a) Muting Global Component



(b) Muting Local Component



(c) Actual Path of Innovation



IA3.3 Public and Private Innovations

Figure IA12 illustrates continental shares in global innovations, separately for public and private innovations. Patterns are similar to those observed for total innovations. However, we detect differential timing in when regions rise to global dominance: Dominance in private innovation activity appears to lead dominance in public innovation activity.

Both private and public innovation activity exhibit substantial persistence at the country level. The autocorrelograms in Figure IA13 display positive autocorrelation coefficients that decay gradually over horizons of several decades, consistent with the view that innovative capacity builds up slowly

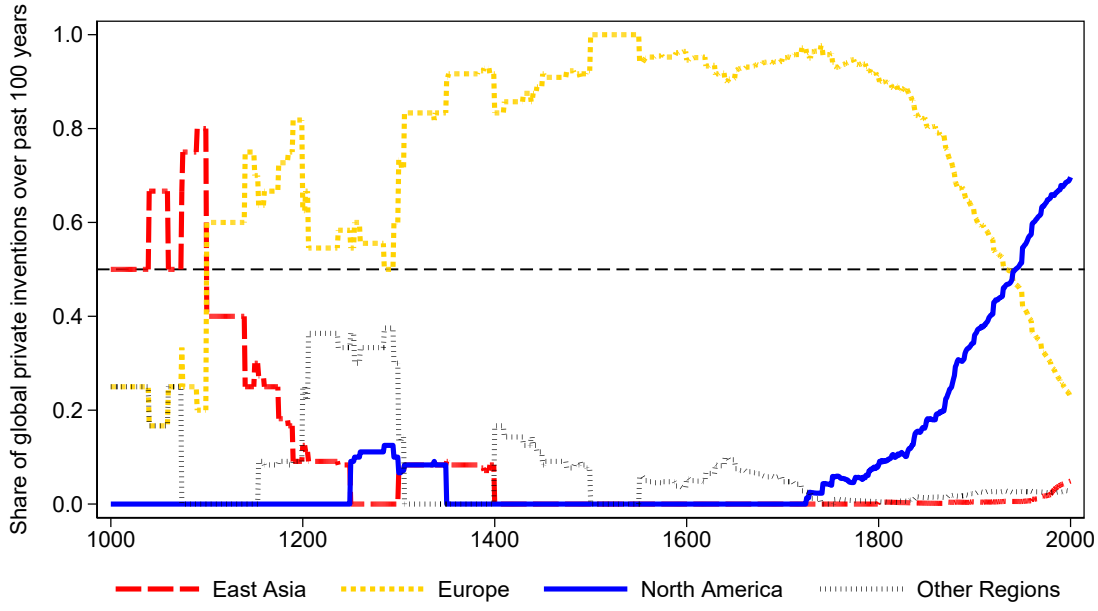
through the accumulation of knowledge, human capital, and specialized infrastructure and dissipates only gradually. The cross-correlograms shed further light on the temporal relationship between the two sectors. Private innovation activity is a positive predictor of public innovation in subsequent decades, with the reverse pattern less pronounced. Notably, the persistence of private innovation is somewhat larger, pointing to private innovative capacity as the more durable component of a country's overall innovation intensity.

Figure IA14 zooms into episodes in which countries become global innovation leaders. We find that both private and public innovation activity rise sharply in the years surrounding the transition to global leadership, with the increase apparent well before the peak and sustained for decades afterward. The rise in private innovation is particularly pronounced. Public innovation also rises during these episodes, but its trajectory is somewhat more diffuse and less sharply timed around the transition.

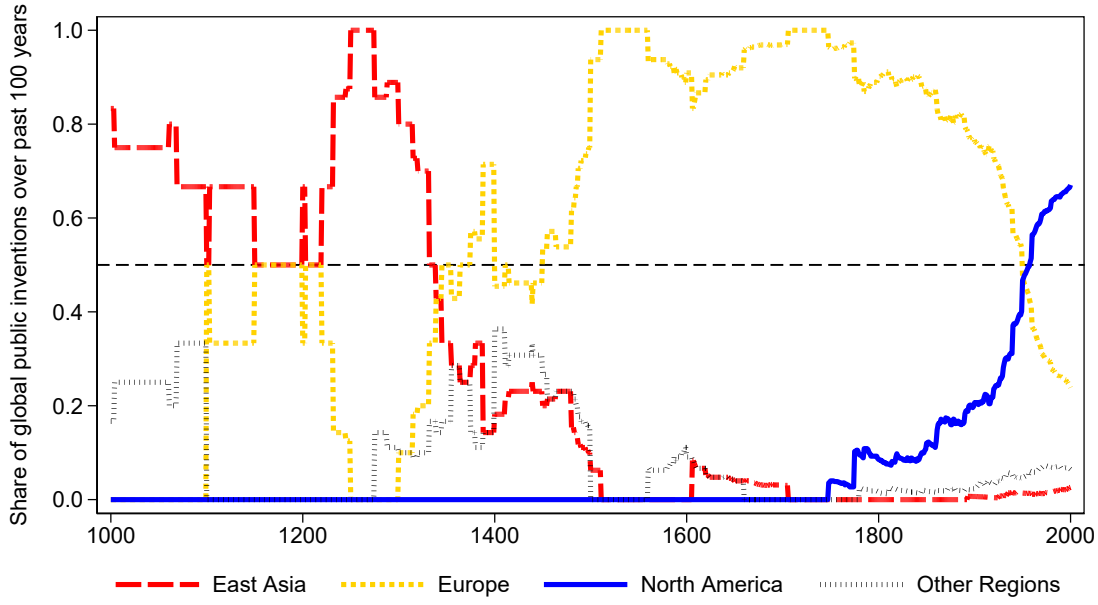
Taken together, the evidence appears inconsistent with the characterization of private innovation activity as a mere byproduct or offshoot of public research endeavors. Government and university activity has tended to accompany rather than precede surges in private innovations.

Figure IA12. Innovation Concentration in Private vs. Public Innovations

(a) Private Innovations



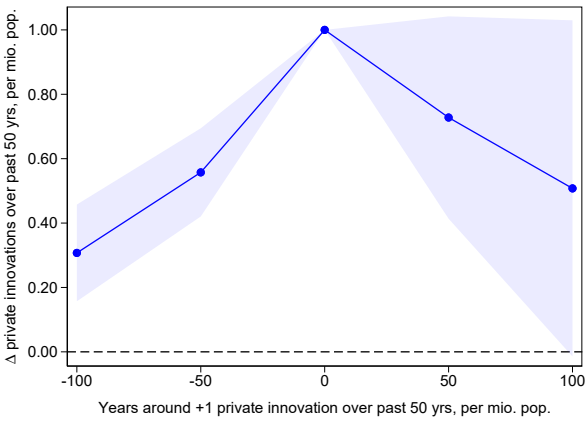
(b) Public Innovations



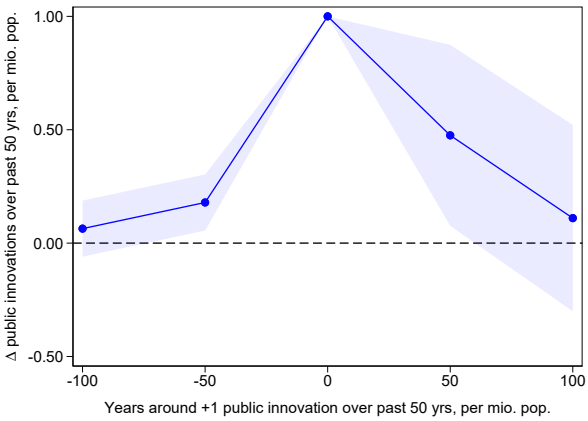
Notes: The figure illustrates the weight of continents in private and public innovations between years 1000 and 2000. Public innovations are defined as any innovations by universities or government entities and innovations classified as military innovations and basic research. Innovations are measured as cumulative sum over the past 100 years, to address lumpiness in the time series in early parts of the sample.

Figure IA13. Dynamic Correlations between Private and Public Innovation Activity

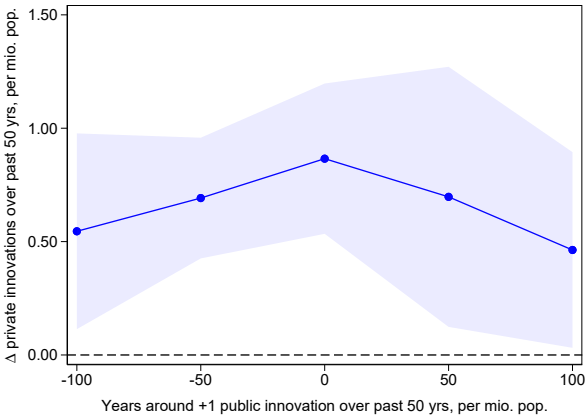
(a) Autocorrelogram Private Innovations



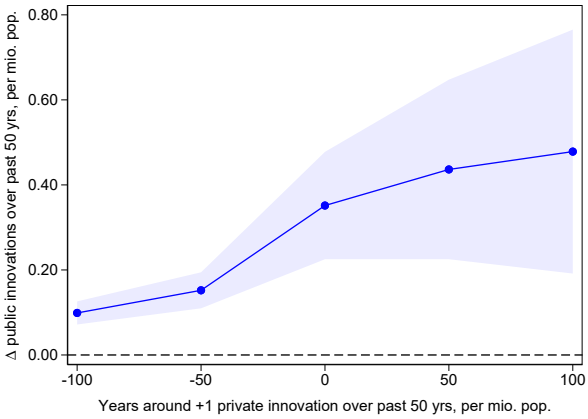
(b) Autocorrelogram Public Innovations



(c) Crosscorrelogram Private-Public



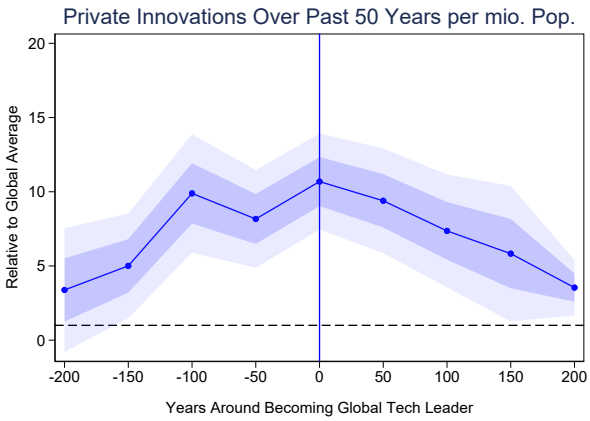
(d) Crosscorrelogram Public-Private



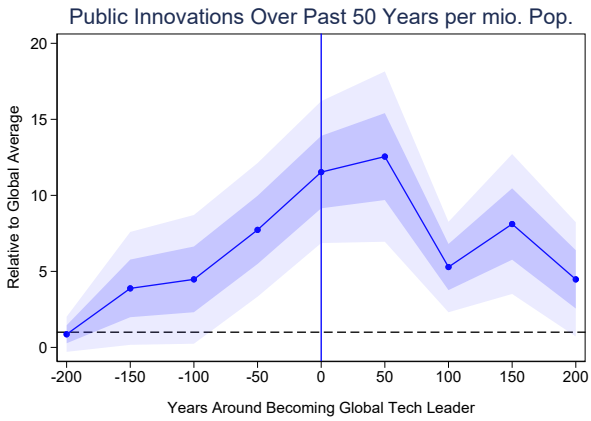
Notes: The figure plots coefficients from regressions of leads and lags of public and private innovation activity at the country level. Innovation activity is measured as the number of public or private innovations over the past 50 years relative to current population in millions. The sample comprises years between 1500 and 2000. Shaded areas mark 95% confidence intervals.

Figure IA14. Private and Public Innovation Activity by Global Innovation Leaders

(a) Private Innovations



(b) Public Innovations



Notes: The figure illustrates innovation activity of global innovation leaders relative to the global average. Countries with the most innovations over the past 50 years per capita are defined to be global innovation leaders. We then study years at which countries became leader and average the number of private and public innovations over the past 50 years per capita at different horizons. We normalize by the prevailing global average to capture general trends in innovation activity over time. Panel (a) shows results for private innovations. Panel (b) shows results for public innovations. Light shaded areas show 95% confidence intervals based on robust standard errors, dark shaded areas show ± 1 standard error.

IA3.4 Decomposing Innovation Concentration

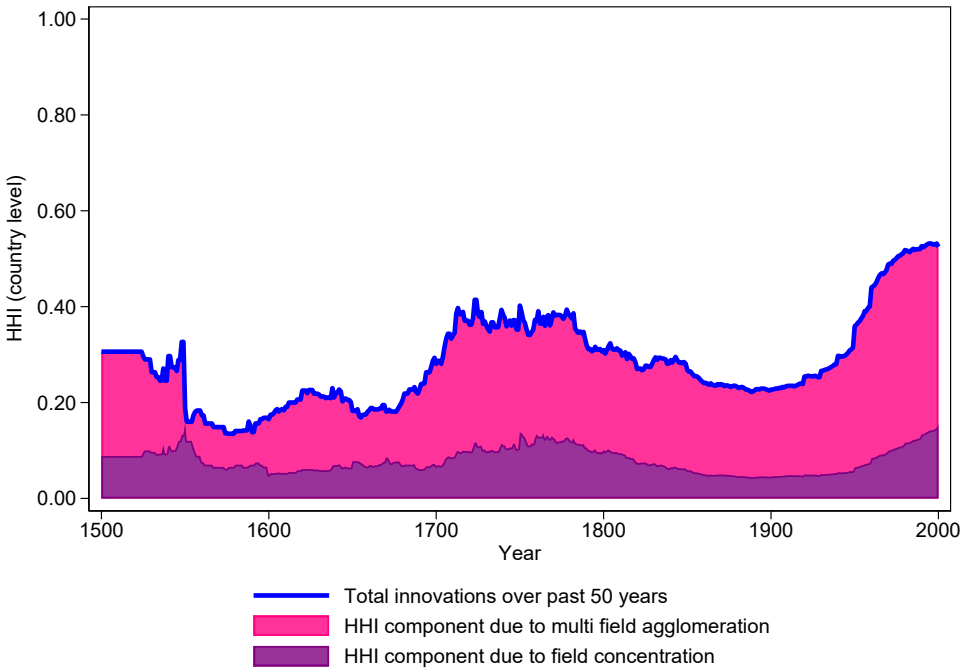
We decompose the Hirschman-Herfindahl index of innovation activity across countries as follows:

$$
\begin{aligned}
HHI_t &= \sum_{c=1}^{C} s_c^2 \\
&= \sum_{c=1}^{C} \left(\frac{\sum_{f=1}^{F} I_{f,c}}{I} \right)^2 \\
&= \sum_{c=1}^{C} \left\{ \left(\frac{\sum_{f=2}^{F} I_{f,c}}{I} \right)^2 + 2 \frac{\sum_{f=2}^{F} I_{f,c}}{I} \frac{I_{1,c}}{I} + \left(\frac{I_{1,c}}{I} \right)^2 \right\} \\
&= 2 \sum_{c=1}^{C} \sum_{f}^{F-1} \sum_{g=f+1}^{F} \frac{I_{f,c}}{I} \frac{I_{g,c}}{I} + \sum_{c=1}^{C} \sum_{f} \left(\frac{I_{f,c}}{I} \right)^2 \\
&= 2 \sum_{c=1}^{C} \sum_{f}^{F-1} \sum_{g=f+1}^{F} \frac{I_{f,c}}{I} \frac{I_{g,c}}{I} + \sum_{f} \left(\frac{I_f}{I} \right)^2 \sum_{c=1}^{C} \left(\frac{I_{f,c}}{I_f} \right)^2 \\
&= 2 \sum_{c=1}^{C} \sum_{f}^{F-1} \sum_{g=f+1}^{F} \frac{I_{f,c}}{I} \frac{I_{g,c}}{I} + \sum_{f=1}^{F} \left(\frac{I_f}{I} \right)^2 HHI_{f,t}
\end{aligned}
$$

The decomposition isolates two conceptually distinct sources of aggregate innovation concentration. The first term—represented in pink in Figure IA15—measures the co-movement of innovation activity across different technological fields within the same country. This term is large when a country that is a strong innovator in one field is simultaneously a strong innovator in others, capturing the tendency of innovative capacity to agglomerate broadly across domains within a single location. The second term—represented in purple—is a weighted average of field-specific HHIs, with weights equal to each field's share of total global innovations squared. It captures within-field concentration: the degree to which each technological domain is itself dominated by a small number of countries. Together, the two terms add up to the overall country-level innovation HHI, offering a transparent decomposition of aggregate concentration into a cross-field agglomeration component and a within-field dominance component.

Figure IA15 shows that both components contribute meaningfully to aggregate concentration throughout the sample. The weighted average of field-specific HHIs accounts for a considerable but largely stable portion of the overall HHI at each point in time. By contrast, the cross-field co-movement term appears to be largely responsible for major swings and the secular increase in innovation activity across the centuries. This indicates that globally leading innovators tend to be strong across multiple fields simultaneously rather than specializing narrowly in a single domain—and increasingly so over time.

Figure IA15. Decomposition of Innovation Activity HHI



Notes: The figure plots the HHI of countries' innovation activity in blue alongside a decomposition into weighted field-specific HHIs and a measure of spatial agglomeration of innovation across fields. Innovation activity is measured as the number of innovations over the past 50 years. The decomposition is: $HHI_t = 2\sum_{c=1}^{C}\sum_{f}^{F-1}\sum_{g=f+1}^{F}\frac{I_{f,c}}{I}\frac{I_{g,c}}{I} + \sum_{f=1}^{F}\left(\frac{I_f}{I}\right)^2 HHI_{f,t}$ where t indexes years, c indexes countries, f indexes fields and $HHI_{f,t}$ is the HHI of country-level innovations in field f in year t. The pink area represents the first term, capturing the co-movement in innovation activity across field within countries. The purple area represents the second term, capturing some weighted average of field-specific concentration.

IA4 Proofs

IA4.1 Royalties

Define $F_{\ell t}^{new}(q_1, q_2)$ as

$$F_{\ell t}^{new}(q_1, q_2) = e^{(\lambda_{\ell t}-\lambda_{\ell t-1})q_2^{-\theta}}\left[1 + (\lambda_{\ell t} - \lambda_{\ell t-1})\left(q_2^{-\theta} - q_1^{-\theta}\right)\right]$$

to be the joint distribution of best and second best techniques within a domain among techniques discovered in ℓ at time t. This will be an input into the royalty payments.

Claim 1. *Total royalty payments from country ℓ_0 paid to researchers in ℓ is*

$$T_{\ell\ell_0 t} = \eta_{\ell\ell_0 t} Y_{\ell_0 t} P_{\ell_0 t} \frac{\varepsilon - 1}{\theta} \frac{\mu_{\ell_0 t} - 1}{\mu_{\ell_0 t}} \frac{(\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta}}{\sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t} \tau_{\ell\tilde{\ell},t}^{-\theta}}$$

Proof. The probability that the best idea ℓ for j is from ℓ', was invented contemporaneously, has productivity q_1, and the second best idea has productivity less than q_2 is

$$\begin{aligned}
& \frac{1}{\tau_{\ell\ell_0}} \frac{\partial F_{\ell t}^{new}\left(\frac{q_1}{\tau_{\ell\ell_0}}, \frac{q_2}{\tau_{\ell\ell_0}}\right)}{\partial q_1} F_{\ell t}^{old}\left(\frac{q_1}{\tau_{\ell\ell_0}}, \frac{q_2}{\tau_{\ell\ell_0}}\right) \prod_{\ell' \neq \ell} F_{\ell t}\left(\frac{q_2}{\tau_{\ell'\ell_0}}\right) \\
= \;& e^{(\lambda_{\ell t} - \lambda_{\ell t-1}) q_2^{-\theta}} (\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta} \theta q_1^{-\theta-1} e^{-\lambda_{\ell t-1} \tau_{\ell\ell_0}^{-\theta} q_2^{-\theta}} \prod_{\ell' \neq \ell} e^{-\lambda_{\ell' t} \tau_{\ell'\ell_0}^{-\theta} q_2^{-\theta}} \\
= \;& (\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta} \theta q_1^{-\theta-1} \prod_{\ell'} e^{-\lambda_{\ell' t} \tau_{\ell'\ell_0}^{-\theta} q_2^{-\theta}} \\
= \;& (\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta} \theta q_1^{-\theta-1} e^{-\Phi_{\ell_0 t} q_2^{-\theta}}
\end{aligned}$$

where

$$\Phi_{\ell_0 t} \equiv \sum_{\ell'} \lambda_{\ell' t} \tau_{\ell'\ell_0}^{-\theta}$$

The profit (before any royalty payments) an entrepreneur in ℓ would earn when producing with productivity q is

$$Y_{\ell t} P_{\ell t}^{\varepsilon} \left(p_{j\ell t}^{1-\varepsilon}\right) \frac{\mu_{\ell t} - 1}{\mu_{\ell t}} = Y_{\ell t} P_{\ell t}^{\varepsilon} \left(\mu_{\ell t} w_{\ell t}^{Y}\right)^{1-\varepsilon} \frac{\mu_{\ell t} - 1}{\mu_{\ell t}} q^{\varepsilon - 1}$$

Thus total payments to inventors in ℓ by entrepreneurs in ℓ_0 is

$$\begin{aligned}
& \int_0^{\infty} \int_0^{q_1} Y_{\ell_0 t} P_{\ell_0 t}^{\varepsilon} \left(\mu_{\ell_0 t} w_{\ell_0 t}^{Y}\right)^{1-\varepsilon} \frac{\mu_{\ell_0 t} - 1}{\mu_{\ell_0 t}} \left[q_1^{\varepsilon-1} - q_2^{\varepsilon-1}\right] \\
& (\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta} \theta q_1^{-\theta-1} \Phi_{\ell_0 t} \theta q_2^{-\theta-1} e^{-\Phi_{\ell_0 t} q_2^{-\theta}} dq_2 dq_1 \\
= \;& Y_{\ell_0 t} P_{\ell_0 t}^{\varepsilon} \left(\mu_{\ell_0 t} w_{\ell_0 t}^{Y}\right)^{1-\varepsilon} \frac{\mu_{\ell_0 t} - 1}{\mu_{\ell_0 t}} (\lambda_{\ell t} - \lambda_{\ell t-1})\, \tau_{\ell\ell_0}^{-\theta} \\
& \int_0^{\infty} \int_0^{q_1} \left[q_1^{\varepsilon-1} - q_2^{\varepsilon-1}\right] \theta q_1^{-\theta-1} \Phi_{\ell_0 t} \theta q_2^{-\theta-1} e^{-\Phi_{\ell_0 t} q_2^{-\theta}} dq_2 dq_1
\end{aligned}$$

The integral has two terms and we solve for them one at a time

$$\begin{aligned}
&\int_0^\infty \int_0^{q_1} q_1^{\varepsilon-1}\theta q_1^{-\theta-1}\Phi_{\ell_0 t}\theta q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 dq_1 \\
= &\int_0^\infty q_1^{\varepsilon-1}\theta q_1^{-\theta-1}e^{-\Phi_{\ell_0 t}q_1^{-\theta}}dq_1 \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\int_0^\infty \left(\Phi_{\ell_0 t}q_1^{-\theta}\right)^{-\frac{\varepsilon-1}{\theta}}\theta q_1^{-\theta-1}e^{-\Phi_{\ell_0 t}q_1^{-\theta}}dq_1 \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)
\end{aligned}$$

$$\begin{aligned}
&\int_0^\infty \int_0^{q_1} q_2^{\varepsilon-1}\theta q_1^{-\theta-1}\Phi_{\ell_0 t}\theta q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 dq_1 \\
= &\int_0^\infty \int_{q_2}^\infty q_2^{\varepsilon-1}\theta q_1^{-\theta-1}\Phi_{\ell_0 t}\theta q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_1 dq_2 \\
= &\int_0^\infty q_2^{\varepsilon-1}q_2^{-\theta}\Phi_{\ell_0 t}\theta q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 \\
= &\int_0^\infty \left(q_2^{-\theta}\right)^{1-\frac{\varepsilon-1}{\theta}}\theta\Phi_{\ell_0 t}q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\int_0^\infty \left(\Phi_{\ell_0 t}q_2^{-\theta}\right)^{1-\frac{\varepsilon-1}{\theta}}\theta\Phi_{\ell_0 t}q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\Gamma\left(2-\frac{\varepsilon-1}{\theta}\right) \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\left(1-\frac{\varepsilon-1}{\theta}\right)\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)
\end{aligned}$$

Together, the integral is

$$\begin{aligned}
&\int_0^\infty \int_0^{q_1} \left[q_1^{\varepsilon-1}-q_2^{\varepsilon-1}\right]\theta q_1^{-\theta-1}\Phi_{\ell_0 t}\theta q_2^{-\theta-1}e^{-\Phi_{\ell_0 t}q_2^{-\theta}}dq_2 dq_1 \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)-\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\left(1-\frac{\varepsilon-1}{\theta}\right)\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right) \\
= &\Phi_{\ell_0 t}^{\frac{\varepsilon-1}{\theta}-1}\frac{\varepsilon-1}{\theta}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)
\end{aligned}$$

So the payment (across all varieties) to contemporaneous innovators in ℓ is

$$T_{\ell\ell_0 t}=N_{\ell t}\eta_{\ell\ell_0 t}\frac{\varepsilon-1}{\theta}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)Y_{\ell_0 t}P_{\ell_0 t}^{\varepsilon}\left(\mu_{\ell_0 t}w_{\ell_0 t}^{Y}\right)^{1-\varepsilon}\frac{\mu_{\ell_0 t}-1}{\mu_{\ell_0 t}}\left(\lambda_{\ell t}-\lambda_{\ell t-1}\right)\tau_{\ell\ell_0}^{-\theta}\frac{1}{\Phi_{\ell_0 t}^{1-\frac{\varepsilon-1}{\theta}}}$$

The price level is $P_{\ell t}=\left(\int_0^{N_{\ell t}}p_{j\ell t}^{1-\varepsilon}dj\right)^{\frac{1}{1-\varepsilon}}=\left(\int_0^{N_{\ell t}}\left(\mu_{\ell t}\frac{w_{\ell t}^{Y}}{q_{j\ell t}}\right)^{1-\varepsilon}dj\right)^{\frac{1}{1-\varepsilon}}=\frac{N_{\ell t}^{\frac{1}{1-\varepsilon}}\mu_{\ell t}w_{\ell t}^{Y}}{Q_{\ell t}}=\frac{N_{\ell t}^{\frac{1}{1-\varepsilon}}\mu_{\ell t}w_{\ell t}^{Y}}{\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{\frac{1}{\varepsilon-1}}\left(\sum_{\tilde{\ell}}\lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}\right)^{\frac{1}{\theta}}}$,

so the payment is

$$\begin{aligned} T_{\ell\ell_0 t} &= N_{\ell t}\eta_{\ell\ell_0 t}\frac{\varepsilon-1}{\theta}\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)Y_{\ell_0 t}P_{\ell_0 t}\left(\frac{N_{\ell t}^{\frac{1}{1-\varepsilon}}\mu_{\ell t}w_{\ell t}^{Y}}{\Gamma\left(1-\frac{\varepsilon-1}{\theta}\right)^{\frac{1}{\varepsilon-1}}\left(\sum_{\tilde{\ell}}\lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}\right)^{\frac{1}{\theta}}}\right)^{\varepsilon-1} \\ &\quad\times\left(\mu_{\ell_0 t}w_{\ell_0 t}\right)^{1-\varepsilon}\frac{\mu_{\ell_0 t}-1}{\mu_{\ell_0 t}}\left(\lambda_{\ell t}-\lambda_{\ell t-1}\right)\tau_{\ell\ell_0}^{-\theta}\frac{1}{\Phi_{\ell_0 t}^{1-\frac{\varepsilon-1}{\theta}}} \\ &= \eta_{\ell\ell_0 t}Y_{\ell_0 t}P_{\ell_0 t}\frac{\varepsilon-1}{\theta}\frac{\mu_{\ell_0 t}-1}{\mu_{\ell_0 t}}\frac{\left(\lambda_{\ell t}-\lambda_{\ell t-1}\right)\tau_{\ell\ell_0}^{-\theta}}{\Phi_{\ell_0 t}} \end{aligned}$$

□

IA4.2 Log linearization

The three equations are

$$\hat{\lambda}_{\ell t}-\frac{\hat{\lambda}_{\ell t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}=s_{\ell t}^{R}\left(s_{\ell t-1}^{R}\right)^{\zeta}\hat{L}_{\ell t}\hat{L}_{\ell t-1}^{\zeta}Z_{\ell t}\Gamma\left(1-\beta\right)\left(\sum_{\tilde{\ell}}\frac{\hat{\lambda}_{\tilde{\ell}t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\kappa_{\ell\tilde{\ell},t}^{-\theta}\right)^{\beta}$$

$$\phi_{\ell t}\equiv\eta_{\ell t}\frac{\varepsilon-1}{\theta}\frac{\left(\lambda_{\ell t}-\lambda_{\ell t-1}\right)}{\sum_{\tilde{\ell}}\lambda_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}=\eta_{\ell t}\frac{\varepsilon-1}{\theta}\frac{\hat{\lambda}_{\ell t}-\frac{\hat{\lambda}_{\ell t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}}{\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell},t}\tau_{\ell\tilde{\ell},t}^{-\theta}}$$

$$s_{\ell t}^{R}=\frac{\left[\phi_{\ell t}B_{\ell t}^{R}\right]^{\frac{\sigma}{1+\sigma}}}{\left[\phi_{\ell t}B_{\ell t}^{R}\right]^{\frac{\sigma}{1+\sigma}}+\left[\left(1-\phi_{\ell t}\right)B_{\ell t}^{A}\right]^{\frac{\sigma}{1+\sigma}}+\left(\frac{1}{\mu_{\ell t}-1}B_{\ell t}^{Y}\right)^{\frac{\sigma}{1+\sigma}}}$$

Suppose that $\kappa_{\ell\tilde{\ell},t}$, $\tau_{\ell\tilde{\ell},t}$, and $\mu_{\ell t}$ are constant over time and $\zeta=0$. Letting a tilde denote the log deviation from steady state, we have

$$\begin{aligned} \frac{\hat{\lambda}_{\ell t}-\frac{1}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\hat{\lambda}_{\ell t-1}}{\hat{\lambda}_{\ell}^{ss}-\frac{1}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\hat{\lambda}_{\ell}^{ss}} &= \frac{s_{\ell t}^{R}}{s_{\ell}^{R,ss}}\frac{\hat{L}_{\ell t}}{\hat{L}_{\ell}^{ss}}\frac{Z_{\ell t}}{Z_{\ell}^{ss}}\left(\frac{\sum_{\tilde{\ell}}\frac{\hat{\lambda}_{\tilde{\ell}t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}}\frac{\hat{\lambda}_{\tilde{\ell}}^{ss}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}\kappa_{\ell\tilde{\ell}}^{-\theta}}\right)^{\beta} \\ &= \frac{s_{\ell t}^{R}}{s_{\ell}^{R,ss}}\frac{\hat{L}_{\ell t}}{\hat{L}_{\ell}^{ss}}\frac{Z_{\ell t}}{Z_{\ell}^{ss}}\left(\frac{\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell}}^{ss}\kappa_{\ell\tilde{\ell}}^{-\theta}\frac{\hat{\lambda}_{\tilde{\ell}t-1}}{\hat{\lambda}_{\tilde{\ell}}^{ss}}}{\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell}}^{ss}\kappa_{\ell\tilde{\ell}}^{-\theta}}\right)^{\beta} \end{aligned}$$

$$\frac{\tilde{\lambda}_{\ell t}-(1+\gamma)^{-\frac{1+\zeta}{1-\beta}}\tilde{\lambda}_{\ell t-1}}{1-(1+\gamma)^{-\frac{1+\zeta}{1-\beta}}}=\tilde{s}_{\ell t}^{R}+\tilde{L}_{\ell t}+\tilde{Z}_{\ell t}+\beta\sum_{\tilde{\ell}}\frac{\hat{\lambda}_{\tilde{\ell}}^{ss}\kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}}\hat{\lambda}_{\tilde{\ell}}^{ss}\kappa_{\ell\tilde{\ell}}^{-\theta}}\tilde{\lambda}_{\tilde{\ell}t-1}$$

Define $x_{\ell t} \equiv \frac{\tilde{\lambda}_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \tilde{\lambda}_{\ell t-1}}{1-(1+\gamma)^{-\frac{1+\zeta}{1-\beta}}}$ to be the innovation in time t, this is

$$x_{\ell t} = \tilde{s}^R_{\ell t} + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} + \beta \sum_{\tilde{\ell}} \frac{\hat{\lambda}^{ss}_{\tilde{\ell}} \kappa^{-\theta}_{\ell\tilde{\ell}}}{\sum_{\tilde{\ell}} \hat{\lambda}^{ss}_{\tilde{\ell}} \kappa^{-\theta}_{\ell\tilde{\ell}}} \tilde{\lambda}_{\tilde{\ell} t-1}$$

The labor equation is

$$\begin{aligned}
s^R_{\ell t} &= \frac{\left[\phi_{\ell t} B^R_{\ell t}\right]^{\frac{\sigma}{1+\sigma}}}{\left[\phi_{\ell t} B^R_{\ell t}\right]^{\frac{\sigma}{1+\sigma}} + \left[(1-\phi_{\ell t}) B^A_{\ell t}\right]^{\frac{\sigma}{1+\sigma}} + \left(\frac{1}{\mu_\ell - 1} B^Y_{\ell t}\right)^{\frac{\sigma}{1+\sigma}}} \\
\frac{s^R_{\ell t}}{s^{R,ss}_\ell} &= \frac{\left[\frac{\phi_{\ell t}}{\phi^{ss}_\ell} \frac{B^R_{\ell t}}{B^{R,ss}_\ell}\right]^{\frac{\sigma}{1+\sigma}}}{s^{R,ss}_\ell \left[\frac{\phi_{\ell t}}{\phi^{ss}_\ell} \frac{B^R_{\ell t}}{B^{R,ss}_\ell}\right]^{\frac{\sigma}{1+\sigma}} + s^{A,ss}_\ell \left[\frac{1-\phi_{\ell t}}{1-\phi^{ss}_\ell} \frac{B^A_{\ell t}}{B^{A,ss}_\ell}\right]^{\frac{\sigma}{1+\sigma}} + s^{Y,ss}_\ell \left(\frac{B^Y_{\ell t}}{B^{Y,ss}_\ell}\right)^{\frac{\sigma}{1+\sigma}}}
\end{aligned}$$

or

$$\begin{aligned}
\tilde{s}^R_{\ell t} &= \frac{\sigma}{1+\sigma}\left(\tilde{\phi}_{\ell t} + \tilde{B}^R_{\ell t}\right) - \frac{\sigma}{1+\sigma}\left[s^{R,ss}_\ell \left(\tilde{\phi}_{\ell t} + \tilde{B}^R_{\ell t}\right) + s^{A,ss}_\ell \left(-\frac{\phi^{ss}_\ell}{1-\phi^{ss}_\ell}\tilde{\phi}_{\ell t} + \tilde{B}^A_{\ell t}\right) + s^{Y,ss}_\ell \tilde{B}^Y_{\ell t}\right] \\
&= \frac{\sigma}{1+\sigma}\left\{\left(s^{A,ss}_\ell + s^{Y,ss}_\ell\right)\left(\tilde{\phi}_{\ell t} + \tilde{B}^R_{\ell t}\right) - s^{A,ss}_\ell \left(-\frac{\phi^{ss}_\ell}{1-\phi^{ss}_\ell}\tilde{\phi}_{\ell t} + \tilde{B}^A_{\ell t}\right) - s^{Y,ss}_\ell \tilde{B}^Y_{\ell t}\right\} \\
&= \frac{\sigma}{1+\sigma}\left[1 - s^{R,ss}_\ell + s^{A,ss}_\ell \frac{\phi^{ss}_\ell}{1-\phi^{ss}_\ell}\right]\tilde{\phi}_{\ell t} + \frac{\sigma}{1+\sigma}\left[s^{A,ss}_\ell \left(\tilde{B}^R_{\ell t} - \tilde{B}^A_{\ell t}\right) + s^{Y,ss}_\ell \left(\tilde{B}^R_{\ell t} - \tilde{B}^Y_{\ell t}\right)\right]
\end{aligned}$$

Let $A_{1\ell} = \frac{\sigma}{1+\sigma}\left[1 - s^{R,ss}_\ell + s^{A,ss}_\ell \frac{\phi^{ss}_\ell}{1-\phi^{ss}_\ell}\right]$, so this can be expressed as

$$\tilde{s}^R_{\ell t} = A_{1\ell}\tilde{\phi}_{\ell t} + \frac{\sigma}{1+\sigma}\left[s^{A,ss}_\ell \left(\tilde{B}^R_{\ell t} - \tilde{B}^A_{\ell t}\right) + s^{Y,ss}_\ell \left(\tilde{B}^R_{\ell t} - \tilde{B}^Y_{\ell t}\right)\right]$$

Suppose adopters can use foreign technologies only with a lag, so that

$$\begin{aligned}
\phi_{\ell t} &\equiv \eta_{\ell t} \frac{\varepsilon - 1}{\theta} \frac{(\lambda_{\ell t} - \lambda_{\ell t-1})}{(\lambda_{\ell t} - \lambda_{\ell t-1}) + \sum_{\tilde{\ell}} \lambda_{\tilde{\ell},t-1} \tau^{-\theta}_{\ell\tilde{\ell}}} = \eta_{\ell t} \frac{\varepsilon - 1}{\theta} \frac{\hat{\lambda}_{\ell t} - \frac{\hat{\lambda}_{\ell t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}}}{\hat{\lambda}_{\ell t} - \frac{\hat{\lambda}_{\ell t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}} + \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell},t-1}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}} \tau^{-\theta}_{\ell\tilde{\ell}}} \\
\tilde{\phi}_{\ell t} &= \tilde{\eta}_{\ell t} + x_{\ell t} - \frac{\left[\hat{\lambda}^{ss}_\ell - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \hat{\lambda}^{ss}_\ell\right] x_{\ell t} + \sum_{\tilde{\ell}} \frac{\hat{\lambda}^{ss}_{\tilde{\ell}}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}} \tau^{-\theta}_{\ell\tilde{\ell}} \tilde{\lambda}_{\ell,t-1}}{\hat{\lambda}^{ss}_\ell - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \hat{\lambda}^{ss}_\ell + \sum_{\tilde{\ell}} \frac{\hat{\lambda}^{ss}_{\tilde{\ell}}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}}} \tau^{-\theta}_{\ell\tilde{\ell}}}
\end{aligned}$$

Let $A_{2\ell} = \frac{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}{(1+\gamma)^{\frac{1+\zeta}{1-\beta}} \hat{\lambda}_{\ell}^{ss} - \hat{\lambda}_{\ell}^{ss} + \sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}$, so that this is

$$\tilde{\phi}_{\ell t} = \tilde{\eta}_{\ell t} + (1 - A_{2\ell})\, x_{\ell t} - A_{2\ell} \frac{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta} \tilde{\lambda}_{\tilde{\ell},t-1}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}$$

All three equations together give

$$\begin{aligned}
\frac{\tilde{\lambda}_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \tilde{\lambda}_{\ell t-1}}{1 - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}}} &= \tilde{s}_{\ell t}^{R} + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} + \beta \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell} t-1} \\
&= A_{1\ell} \tilde{\phi}_{\ell t} + \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} + \beta \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{s}} \\
&= A_{1\ell} \left\{ \tilde{\eta}_{\ell t} + (1 - A_{2\ell}) \frac{\tilde{\lambda}_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \tilde{\lambda}_{\ell t-1}}{1 - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}}} - A_{2\ell} \frac{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta} \tilde{\lambda}_{\tilde{\ell},t-1}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \right\} + \frac{\sigma}{1+\sigma} \left[s
\end{aligned}$$

Let $x_{\ell t} \equiv \frac{\tilde{\lambda}_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \tilde{\lambda}_{\ell t-1}}{1 - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}}}$, so this is

$$\begin{aligned}
x_{\ell t} &= \tilde{s}_{\ell t}^{R} + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} + \beta \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell} t-1} \\
&= A_{1\ell} \tilde{\phi}_{\ell t} + \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} + \beta \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell} t-1} \\
&= A_{1\ell} \left\{ \tilde{\eta}_{\ell t} + (1 - A_{2\ell})\, x_{\ell t} - A_{2\ell} \frac{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta} \tilde{\lambda}_{\tilde{\ell},t-1}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \right\} + \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell}
\end{aligned}$$

or

$$[1 - A_{1\ell}(1 - A_{2\ell})]\, x_{\ell t} = A_{1\ell} \left\{ \tilde{\eta}_{\ell t} - A_{2\ell} \frac{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta} \tilde{\lambda}_{\tilde{\ell},t-1}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \right\} + \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell t}$$

Or

$$x_{\ell t} = -\frac{A_{2\ell} A_{1\ell}}{1 - A_{1\ell}(1 - A_{2\ell})} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell},t-1} + \frac{\beta}{1 - A_{1\ell}(1 - A_{2\ell})} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{\lambda}_{\tilde{\ell} t-1} + \varsigma_{\ell t}$$

where

$$\varsigma_{\ell t} = \frac{1}{1 - A_{1\ell}(1 - A_{2\ell})} \left\{ A_{1\ell} \tilde{\eta}_{\ell t} + \frac{\sigma}{1+\sigma} \left[s_{\ell}^{A,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{A} \right) + s_{\ell}^{Y,ss} \left(\tilde{B}_{\ell t}^{R} - \tilde{B}_{\ell t}^{Y} \right) \right] + \tilde{L}_{\ell t} + \tilde{Z}_{\ell t} \right\}$$

Taking a quasi first difference gives

$$x_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} x_{\ell t} = -B_{1\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \tau_{\ell\tilde{\ell}}^{-\theta}} \tilde{x}_{\ell,t-1} + B_{2\ell} \sum_{\tilde{\ell}} \frac{\hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}}{\sum_{\tilde{\ell}} \hat{\lambda}_{\tilde{\ell}}^{ss} \kappa_{\ell\tilde{\ell}}^{-\theta}} \tilde{x}_{\tilde{\ell}t-1} + \varepsilon_{\ell t}$$

where

$$\begin{aligned} B_{1\ell} &\equiv \frac{\left[1 - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}}\right] A_{2\ell} A_{1\ell}}{1 - A_{1\ell}(1 - A_{2\ell})} \\ B_{2\ell} &\equiv \frac{\beta \left[1 - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}}\right]}{1 - A_{1\ell}(1 - A_{2\ell})} \end{aligned}$$

$$\varepsilon_{\ell t} = \varsigma_{\ell t} - (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} \varsigma_{\ell t}$$

In special case in which there are N continents each with J locations. Let Suppose that costs of distance are such that $\tau_{\ell\tilde{\ell}} = \tau_{within}$ and $\kappa_{\ell\tilde{\ell}} = \kappa_{within}$ if ℓ and $\tilde{\ell}$ are on the same continent, and $\tau_{\ell\tilde{\ell}} = \tau_{across}$ and $\kappa_{\ell\tilde{\ell}} = \kappa_{across}$ if ℓ and $\tilde{\ell}$ are on different continents. The countries are othersie symmetric, so that $B_{1\ell} = B_1$ and $B_{2\ell} = B_2$. The system can be expressed as

$$x_{njt+1} = a x_{njt} + b \sum_{\tilde{j} \in J_n \setminus \{j\}} x_{n\tilde{j}t} + c \sum_{\tilde{n} \neq n} \sum_{\tilde{j} \in J_{\tilde{n}}} x_{\tilde{j}\tilde{n}t} + \varepsilon_{njt}$$

with

$$\begin{aligned} a &= (1+\gamma)^{-\frac{1+\zeta}{1-\beta}} - B_1 \frac{1}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{1}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \\ b &= -B_1 \frac{\tau_{within}^{-\theta}}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{\kappa_{within}^{-\theta}}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \\ c &= -B_1 \frac{\tau_{across}^{-\theta}}{\sum_{\tilde{\ell}} \tau_{\ell\tilde{\ell}}^{-\theta}} + B_2 \frac{\kappa_{across}^{-\theta}}{\sum_{\tilde{\ell}} \kappa_{\ell\tilde{\ell}}^{-\theta}} \end{aligned}$$

We can summarize the system as

$$x_{t+1} = M x_t + \varepsilon_{t+1}$$

Define the cross-continent average research and average disturbance as

$$\begin{aligned} \bar{x}_{nt} &= \frac{1}{J} \sum_{j \in J_n} x_{njt} \\ \bar{\varepsilon}_{nt} &= \frac{1}{J} \sum_{j \in J_n} \varepsilon_{njt} \end{aligned}$$

and worldwide averages as

$$\begin{aligned}\bar{x}_t &= \frac{1}{N}\sum_n \bar{x}_{nt} \\ \bar{\varepsilon}_t &= \frac{1}{J}\sum_{j\in J_n} \varepsilon_{nt}\end{aligned}$$

Finally, define a country's deviation from its continent average as and the country's deviation of its disturbance from the continent average is

$$\begin{aligned}y_{njt} &= x_{njt} - \bar{x}_{nt} \\ u_{njt} &= \varepsilon_{njt} - \bar{\varepsilon}_{nt}\end{aligned}$$

$$\begin{aligned}y_{nt} &= \bar{x}_{nt} - \bar{x}_t \\ u_{nt} &= \bar{\varepsilon}_{nt} - \bar{\varepsilon}_t\end{aligned}$$

We then have that y_{jt} and x_{nt} follow the following processes

$$y_{njt+1} = (a-b)\, y_{njt} + u_{njt+1}$$

$$y_{nt+1} = [(a-b) + J\,(b-c)]\, y_{nt} + u_{nt+1}$$

To get these, the country law of motion can be expressed as

$$x_{njt+1} = (a-b)\, x_{njt} + bJ\bar{x}_{nt} + c\,(NJ\bar{x}_t - J\bar{x}_{nt}) + \varepsilon_{njt+1}$$

Average over countries in continent n gives

$$\bar{x}_{nt+1} = (a-b)\, \bar{x}_{nt} + bJ\bar{x}_{nt} + c\,(NJ\bar{x}_t - J\bar{x}_{nt}) + \bar{\varepsilon}_{nt+1}$$

Average over continents gives

$$\bar{x}_{t+1} = (a-b)\, \bar{x}_t + bJ\bar{x}_t + c\,(NJ\bar{x}_t - J\bar{x}_t) + \bar{\varepsilon}_{t+1}$$

Subtracting the second from the first gives

$$y_{njt+1} = (a-b)\, y_{njt} + u_{njt+1}$$

Subtracting the third from the second gives

$$\begin{aligned} y_{nt+1} &= (a-b)\, y_{nt} + bJy_{nt} + c\,(-Jy_{nt}) + u_{nt+1} \\ &= [(a-b) + J\,(b-c)]\, y_{nt} + u_{nt+1} \end{aligned}$$