Innovation

Invention, Knowledge Transfer, and Innovation

Technical Appendix

NSB-2022-4

March 08, 2022

Table of Contents

Technical Appendix

Survey and Alternative Data Are Used in This Report

This **technical appendix** supplements the report text with additional information on the data used in the report, including fitness for use and construction of the indicators. Quality considerations for both survey data and alternative data include three domains: utility, objectivity, and integrity (Dworak-Fisher et al. 2020). Identified quality dimensions within utility are relevance, accessibility, timeliness, punctuality, and granularity. Accuracy, reliability, and coherence are the dimensions of objectivity. Integrity consists of scientific integrity, credibility, computer and physical security, and confidentiality.

Invention Indicators

Inventors around the world obtain protection for inventions through national and regional jurisdictions. The U.S. Patent and Trademark Office (USPTO) is the federal agency responsible for handling patent and trademark applications and issuing patents and trademark registrations in the United States. It grants utility patents, plant patents, design patents, and trademarks (USPTO 2015). USPTO patents protect inventions in the U.S. market. All the patent data presented in this report are for utility patents.

Patent registration data include records for the inventor and for the party to whom legal rights are assigned. Patents have inventors (one or more) and grantees, in which the latter become the owners of the intellectual property (IP) covered by the patent. Both types of information are used in this report. The notes for tables and figures specify the approach for each indicator.

Contributors to U.S. Patent Activity

Patent data used throughout this report are based on the administrative records of the legal authorities granting patents, accessed through publicly available databases. USPTO records are accessed through the USPTO's PatentsView database. The National Center for Science and Engineering Statistics (NCSES), SRI International, and Science-Metrix prepared the tabulations from the databases presented in this report. Patent-level data were downloaded from the PatentsView data analysis and visualization platform maintained by the USPTO in collaboration with other federal agencies and academic institutions. In PatentsView, USPTO patent data are parsed and structured into a relational database. PatentsView assigns patents to their relevant technology fields based on different classification schemes, including internationally agreed upon technology fields from the World Intellectual Property Organization (WIPO).

PatentsView applies statistical techniques to match and link inventor names and locations. This matching of names and locations, known as *inventor and location disambiguation*, enables analyses of patterns and trends in patenting activity in the United States and abroad. For the USPTO patents by economic sector (Table SINV-1, for example), patents are shown based on the sector of the owner, which can be different from the inventor. When USPTO patents are shown by geography, this corresponds to the residence of the inventor. Detailed technical documentation for the data collection, curation, fractionization, and tabulation is found in the 2022 report commissioned by SRI International: *Patent and Trademark Indicators for the Science and Engineering Indicators 2022: Technical Documentation* prepared by Science-Metrix (2022).

Importance of Intellectual Property Protection to U.S. Business

The 2018 Business Research and Development Survey (BRDS) is the report's data source on the importance of different types of IP to U.S. businesses. The survey is conducted by the U.S. Census Bureau in collaboration with NCSES. It is a sample survey representing for-profit, publicly or privately held companies with 10 or more employees in the United States in the mining, utilities, construction, manufacturing, wholesale trade, retail trade, or services industries. The sample survey is based on the Business Register, the database of U.S. business establishments and companies that is created by

integrating data from business tax returns with data collected in the Economic Census and other Census Bureau surveys. The 2018 data used from BRDS are found at **https://ncses.nsf.gov/pubs/nsf21312**. In subsequent years following 2018, this survey will be replaced with the Business Enterprise Research and Development Survey (BERD, **https://www.nsf.gov/statistics/srvyberd/**).

Geographical Distribution of Patent Activity in the United States

The underlying data for the county-level presentation of USPTO patents for 2020 are prepared by NCSES, SRI International, and Science-Metrix. Patent data are drawn from the USPTO's PatentsView database, described earlier in the **Contributors to U.S. Patent Activity** section.

Figures presented in the report show patent intensity, defined here as the number of patents issued to inventors residing in the county divided by the total population in that county. For patents awarded to inventors in multiple locations, this report uses fractional counts of patents to avoid double counting. With fractional counts, a county receives partial credit for a patent in proportion to the number of named inventors who reside in that county divided by all named inventors.

The patent documents used in the analysis are drawn from the USPTO's PatentsView database in the late spring of 2021. The county population data are the 2020 vintage of U.S. Census data, located at **https://www.census.gov/newsroom/press-releases/2021/2020-vintage-population-estimates.html**.

The methodology for the county-level patent data in this report builds on two main sources. First, the Patent Technology Monitoring Team within the USPTO prepared and released state, county, and metropolitan regional patent data for several years, with tables covering the years 1998 to 2015 (USPTO 2021). An early challenge in this work was geocoding, or the assignment of geographical information to a dataset. For use in regional economic analysis, Carlino, Chatterjee, and Hunt (2006, 2007) used geocoding to extend the information reported in patent documents, allowing assignment to a location.

Second, for the detailed tables by technology area in this report, patents are classified under WIPO's classification of technology areas. This report uses international patent classification (IPC) reformed codes to prepare the patent data; these codes incorporate changes made with the eighth revision of the WIPO classification in 2006. Patents can be tagged with multiple codes and can fall under multiple WIPO technology areas. This report uses fractional counts of patents to avoid double counting across sectors and technology areas.

The public use files accompanying this report show county patent intensity for 2020 in two different ways: based on the location of the inventor and based on the location of the owner. This allows two different kinds of analysis: one based on where inventors live and one based on where the ownership rights have been assigned, which can be a workplace or a residence of the patent owner. The figures in this report are based on the ownership file. These files contain U.S. county-level data for total patents, patents by technical field, and unassigned patents for years 1998–2020 (U.S. Department of Commerce, Census Bureau 2021).

Detailed technical documentation for the preparation and evaluation of these prototype statistics is found in the 2022 report commissioned by SRI International: *Patent and Trademark Indicators for the Science and Engineering Indicators 2022: Technical Documentation* prepared by Science-Metrix (2022).

International Patent Families

The data described in the report as "international patents" measure an original invention and its subsequent extensions as a "family" or group of related patents. A patent family refers to a group of related patents that share a single original invention in common. All subsequent patents in a family refer to the first patent filed, called the priority patent. The year the priority patent is granted is the year when the patent family is counted. The source data for international patent family data for this report is the European Patent Office's (EPO) PATSTAT (**https://www.epo.org/searching-for-patents/business/patstat.html#tab-1**).

As an indicator, patent families provide a broad unduplicated measure of global invention. Patenting standards vary with the jurisdiction. According to the international patent documentation, PATSTAT assigns patent families to geographic locations based on patent inventorship information found on the priority patent. As is done with USPTO patents, PATSTAT allocates patent families among regions, countries, or economies using fractional counts based on the residences of all named inventors.

Using these PATSTAT data from the EPO, international patent families are tabulated by building on a methodology proposed by a team of researchers from academia and the Organisation for Economic Co-operation and Development (OECD). This method uses information within patent families to fill in gaps regarding inventorship for patent offices where data are not complete, looking at related patents in other offices when information is not available for a patent. When information on inventorship cannot be retrieved from any office, the approach relies on the assumption that inventors are frequently from the same country as the assignees who requested the patent, using all available patents within the family to fill remaining gaps. As a final step, for the remaining priority patents with missing information, the country of the patent authority is projected as the country of inventorship, because in most cases patents without any information and no related patent at the world level will be the fruit of local inventors. Overall, the level of precision is high when dealing with large-scale analyses such as the one prepared for this project.

One limitation of the method described above that emerged is that some valid patents are overlooked, leading to an undercount for some countries, including the United States. A published patent may mention an earlier priority document that is not existent in the EPO dataset. Such a document may be missing if the office where it was filed has not published it or if the priority document is not a patent of invention. Artificial priority patents contain only scarce information, including patent office, date, and type of applied document. Critical missing information include names, addresses of applicants and inventors, and IPC codes. A large share of artificial patents come from provisional patents, which are patent applications often used in the United States and other markets to quickly protect an invention at low cost, in the hope of later filing a patent application for a utility patent in the same market. These missing patents account for about 15% of the 2020 patent families.

To address the limitation related to artificial priority patents, instead of dropping patent families that first came to life through these artificial priority patents, the first utility patent in the family that was applied for after the artificial priority patent was selected as the replacement to act as the first priority patent in these cases. This then makes obtainable the relevant information for IPC codes and inventorship in the same manner as is done for other priority patents, following the approach designed by de Rassenfosse et. al (2013). The implementation of this correction leads to more complete information for a large number of countries around the globe, impacting most those that have a national patent office allowing for provisional patent applications, or where inventors often reach out to markets where these are available.

Women as Inventors on Patents

Both the USPTO and WIPO have produced analysis on the role of women as inventors on patents. The indicator shown in Figure INV-11 is the ratio of the number of patent applications with at least one woman listed as an inventor. Names on the patent documents are matched to a database matching names to sex based on a 6.2 million name database for 182 countries, created using country-level sets of name dictionaries. This database is compiled primarily from WIPO's annual IP statistics surveys (see link below) and data compiled by WIPO in processing international applications and registrations through the Patent Cooperation Treaty, Madrid System, and Hague System. The WIPO working paper and the dataset are available at **https://www.wipo.int/ipstats/en/**.

A limitation of this ratio is that it is affected by both the composition of the named inventors but also by the size of the team. A single woman on a large team will count as a patent with a woman inventor. The women inventor rate shows the number of women named as patented inventors in a given year, divided by the total number of inventors.

A primary limitation of either approach is matching of names to either women or men, as this information is not generally captured in the patent documents. Identifying the accuracy of this matching method requires a training set in which sex has already been identified.

Knowledge Transfer Indicators

Matching Citations to Nonpatent Literature

Patent applications filed with the USPTO include citations to other patents. These citations show how a novel invention builds on and distinguishes itself from other patents within the existing technological ecosystem. Some patent applications also include citations to nonpatent literature (NPL), such as peer-reviewed publications. NPL citations show how knowledge flows into inventions. Matching these citations to peer-reviewed scientific publications helps assess the uptake of research in subsequent development efforts.

Science-Metrix matched the NPL citations from PatentsView to records in Scopus, Elsevier's abstract and citation database. An algorithm extracted and parsed the publication titles, years, author names, abbreviated names, volume and issue numbers, and page ranges of research journals and conference proceedings found in NPL citations. Science-Metrix then used statistical techniques to compare these extracted data with information extracted from the Scopus database to match NPL citations in PatentsView to cited publications appearing in Scopus.

University Technology Transfer: AUTM Survey

The source of several of the university technology transfer indicators in this report is the AUTM survey. These data address a policy-relevant set of questions because the Bayh-Dole Act of 1980 (Patent and Trademark Act Amendments of 1980, P.L. 96–517) created a uniform patent policy among the many federal agencies that fund university research, allowing those institutions to retain ownership of inventions made under federally funded research programs. It has been widely regarded as having been an important stimulant for academic institutions to pursue technology transfer activities. The primary federal survey, Higher Education Research and Development (HERD, found at **https://www.nsf.gov/statistics/srvyherd/**) focuses on research and development expenditures rather than technology transfer.

In contrast, the AUTM survey provides data about the changes taking place in university technology transfer since the implementation of Bayh-Dole. AUTM is a membership-based organization for university technology transfer professionals; they survey their members annually on invention, patenting, licensing, and other technology transfer activities. The survey data are collected in AUTM's STATT database, which is available to members and for a subscription fee. In 2019, AUTM had 312 members. Science and Engineering Indicators (SEI) reports have reported these AUTM survey data for several cycles as representative of academic technology transfer.

As a measure of all academic technology transfer, the AUTM data appear to undercount at least some aspects of technology transfer activity. The AUTM survey reports patents granted based on university technology, reporting 5,204 patents issued in 2019 and 5,704 in 2020. Based on the analysis of USPTO utility patents described in this report, the count of academic patents was 7,781 in 2019 and 7,834 in 2020.

In this context, the response rate to the AUTM survey has fallen over the course of several cycles. Correspondence with AUTM reports a response rate for 2019 of 57.9%, or 179 out of 312 member institutions. The response rate of the survey in 2017 was 61.9%. As a result, more of the covered population was imputed in 2019 compared to 2017. Data are subject to revision; however, data from more than two years prior are not updated and are considered part of the historical record (AUTM 2021).

Example of the use of AUTM data:

Aksoy AY, Beaudry C. 2021. How Are Companies Paying for University Research Licenses? Empirical Evidence from University-Firm Technology Transfer. *The Journal of Technology Transfer*. Available at ***https://doi.org/10.1007/s10961-020-09838-x***. Accessed October 2021.

Federal Technology Transfer Annual Reports to Congress

Federal policy supports the transfer of federally owned or originated technology to state and local governments as well as to the private sector and requires that this activity be reported to the president and Congress on an annual basis (Title 15 of U.S. Code, Section 3710(g)(2)). The Stevenson-Wydler Technology and Innovation Act of 1980 (P.L. 96–480) directed federal agencies with laboratory operations to become active in the technology transfer process. It also required these agencies to establish technology transfer offices (termed Offices of Research and Technology Applications) to assist in identifying transfer opportunities and establishing appropriate arrangements for transfer relationships with nonfederal parties.

This statutory report is prepared by the National Institute of Standards and Technology (NIST) and incorporates reporting from federal agencies. The compilation of annual data by NIST for annual reporting purposes provides cross-category comparability of reported indicators by agencies. However, the data are released with a substantial lag, and individual agencies may report more recent or more comprehensive data. The 2016 fiscal year report was released in September 2019 (NIST 2019); annual data are accessed in the Federal Lab Technology Transfer Database v.2015 (last accessed July 2021).

Small Business Innovation Research and Small Business Technology Transfer Metrics

The U.S. Small Business Administration coordinates and helps implement the Small Business Innovation Research (SBIR) and Small Business Technology Transfer programs. Data for this report are released by the Small Business Administration. The dataset includes annual award counts, firm counts, and the amount of the award or obligation (for years before 2015). Data dimensions include participating agency and program phase, Phase 1 and Phase 2. Phase 1 grants are intended to establish the potential viability of a project, while Phase 2 grants focus on continuing the research and development (R&D) activities initiated in Phase 1. The database is updated continually by fiscal year and described and accessed from the SBIR website: **https://www.sbir.gov/analytics-dashboard**.

Example of analytic work with SBIR data:

Audretsch DB, Link AN, Scott JT. 2002. Public/Private Technology Partnerships: Evaluating SBIR-Supported Research. *Research Policy* 31(1)145–158.

Citizen Science Data

This topic is new in SEI 2022. The federal agency data presented in this report are the part of this activity organized and presented by the U.S. government at **www.citizenscience.gov**.

The activity is defined there as "a form of open collaboration in which individuals or organizations participate voluntarily in the scientific process in various ways, including enabling the formulation of research questions; creating and refining project design; conducting scientific experiments; collecting and analyzing data; interpreting the results of data; developing technologies and applications; making discoveries; and solving problems."

The extensive academic literature in this area is outside the scope of this report. However, there are grants from the National Science Foundation and other funders for academic projects that have "citizen science" in the title or abstract. These grants are not included here.

Open-Source Software Data

Federal Contributions

The data in Table INV-4 show the number of open-source software (OSS) repositories with contributors affiliated with the federal government, identified by an email address. The federal agency website **https://code.gov** is the initial source of agency contributors and projects. Most of the projects listed on code.gov provide links to the repository to download OSS code. The applications programming interface (API) provided by code.gov was used to collect data on projects associated with 26 federal government agencies, including links to the code repository. These repositories reside in well-known platforms, such as GitHub, SourceForge, and Bitbucket, as well as web page repositories run by units of the federal government, such as the National Aeronautics and Space Administration, Sandia National Laboratories, and others.

To better identify the contributions of U.S. federal government entities, data collected on these projects from GitHub, currently the world's largest OSS repository, were integrated with repository-level data accessed through the GHOST.jl (2021) software package. This provided a method to measure the contribution of respective agencies based on numbers of code edits, known as commits, and lines of code. For each repository, the commit data are collected for the base branch that includes information on the lines of code added and deleted by each author along with the timestamp and the associated email, name, and user ID. A concordance file was prepared using *The United States Government Manual*, the A-Z Index of U.S. Government Departments and Agencies (**https://www.usa.gov**), the GitHub crowd-source government entities directory, and code.gov.

A caveat to this matching is that using the authors' emails to identify their affiliation allows the tabulated data to include contributions based on the assumption that they are contributing as part of that organization. If business-related addresses were used, as for government contractors, these commits as well as lines of code would not be counted.

International Collaboration in Open-Source Software

The data on international collaborations in OSS for 2018 have been collected from GitHub. Each collaboration counted refers to contributors committing code to a common repository. These data were collected using the GHOST.jl (2021) package to gather data from the machine detectable, Open-Source Initiative (OSI)−approved licensed repositories (shown in the list at the end of this **technical appendix**).

Contributors were linked to countries using an external dataset from GHTorrent, which includes contributor-level attributes such as self-reported location, affiliation, and email information (Gousios and Spinellis 2012). These data were further cleaned to recode common misspellings, major cities, major academic institutions in each country, and email domains.

Innovation Indicators

USPTO Trademarks

Trademarks, which protect original symbols, are issued by national and regional offices. Trademark data used in this report come from two sources: the USPTO and WIPO. WIPO statistics provide broad coverage on annual trademark applications internationally, including by national income level. Trademarks are classified under the 11th edition of the Nice Classification of goods and services, which classifies trademarks under 34 categories of goods and 11 categories of services (**https://www.wipo.int/classifications/nice/en/**).

In this report, trademarks are assigned to geographic locations based on the country of residence of the trademark holders. To avoid double counting, this report uses fractional counts for trademarks shared by holders in multiple locations; a country receives partial credit for a trademark based on the number of trademark holders who reside in that country divided by all of the trademark holders for the particular trademark. This report also uses fractional counts to assign trademarks to the corresponding categories under the classification.

Unlike the county patent data described in the Invention section of the report, the subnational trademark data presented in this report have been developed using existing techniques, but without an existing benchmark. The primary matching strategy uses zip codes, which allows for a cross-reference of county-level U.S. zip codes to U.S. counties. The USPTO trademark data are more complete than the USPTO patent data, thus the address matching is of higher accuracy.

An example of analytic work using trademarks:

von Graevenitz G, Graham SJH, Myers AF. 2021. Distance (Still) Hampers Diffusion of Innovations. *Regional Studies.* Available at **https://www.tandfonline.com/doi/full/10.1080/00343404.2021.1918334**. Accessed October 2021.

PitchBook Venture Capital Data

The venture capital data used in this report supplement the analysis by shedding light on trends in market-driven innovation and entrepreneurship. Venture investors tend to invest in companies and industries with products they believe have a significant likelihood of achieving market success. In this regard, data on U.S. and global venture capital investment trends can be viewed as leading indicators of the innovative output. The data used in Figure INV-23, Figure INV-24, and Figure INV-25 and Table SINV-98 and Table SINV-99 have been accessed from the proprietary database PitchBook. This database has been created and updated by PitchBook Data, Inc., a company that collects financial and business data on the Web and provides subscription-based data (**https://pitchbook.com/**).

Venture capital (VC) presented at the country level is classified based on the location of the company headquarters. The search terms used are Deal Type: All VC Stages; Ownership Status: All Ownership Statuses; Backing Status: VC-backed; Business Status: All Business Statuses.

The classifications shown in Figure INV-24 and Figure INV-25 are PitchBook's industry sectors, classified by PitchBook based on the firm's primary industry. The industries within that classification are described as follows:

- **Information technology**: Communications and networking, computer hardware, semiconductors, information technology services, software, and other information technology.
- **Health care**: Health care devices and supplies, health care services, health care technology systems, pharmaceuticals and biotechnology, and other health care.
- **Business to consumer**: Apparel and accessories, consumer durables, consumer nondurables, retail, transportation, media, restaurants, hotels and leisure, services (nonfinancial), and other consumer product services.
- **Business to business**: Commercial products, commercial services, commercial transportation, and other business products and services.
- **Financial services**: Capital markets/institutions, commercial banks, insurance, and other financial services.
- **Materials and resources:** Agriculture, chemicals and gases, construction (nonwood), forestry, metals, minerals and mining, textiles, and other materials.
- **Energy:** Energy equipment, exploration, production and refining, energy services, utilities, and other energy.

Recent example of the PitchBook database for VC analysis:

Lerner J, Nanda R. 2020. *Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn.* NBER Working Paper 27492. Available at **https://www.nber.org/papers/w27492**. Accessed September 2021.

Business Innovation Survey Data

The *Oslo Manual*, prepared by the OECD and Eurostat, provides a definition for firm-level innovation activity that countries and economies have widely used to enhance comparability of international data (OECD/Eurostat 2005). This framework guides the collection of survey data, including, notably, the Community Innovation Surveys from the European Union Statistical Office and the Business Research and Development and Innovation Survey (BRDIS) from NCSES and the U.S. Census Bureau. Following the *Oslo Manual*, these surveys define innovation as the "implementation of a new or significantly improved product (good or service) or process, a new marketing method, or a new organizational method" (OECD/Eurostat 2005:46–7). The *Oslo Manual* and its definition of innovation were revised in 2018 (OECD/Eurostat 2018). These revisions will guide future surveys and data collection.

Annual Business Survey

Statistics on the introduction of new products and processes by U.S. industries from 2015 to 2017 are from a new set of survey questions introduced into the Annual Business Survey (ABS). Conducted by the U.S. Census Bureau in accordance with an interagency agreement with NCSES in 2019, the data represent an estimated 4.6 million U.S. for-profit companies with one or more employees. Compared to the data used in the *Indicators 2020* "**Invention, Knowledge Transfer, and Innovation**" report, the ABS provides a more comprehensive view of the incidence of innovation by businesses located in the United States. The prior data source, the 2016 BRDIS, asked a more limited set of innovation-related questions representing for-profit firms with five or more employees. As a result of the differences in the two surveys, BRDIS and ABS, the reported U.S. innovation rate for 2014–2016 in BRDIS is lower than that reported for 2015–2018 in ABS.

The ABS is a mandatory, confidential sample survey that collects data on innovation, R&D activity, technology, IP, and business owner characteristics. Firms are identified from the Business Register, the database of U.S. business establishments and companies that is created by integrating data from business tax returns with data collected in the Economic Census and other Census Bureau surveys.

Business Dynamics Statistics

Business Dynamics Statistics (BDS) is a public use dataset prepared and made available by the U.S. Census Bureau's Center for Economic Studies (CES). This report uses data from public-release tables of annual aggregate statistics describing establishment openings and closings, firm startups, and job creation and destruction. CES compiles the database using data from the Economic Census and administrative data. Methodological details about BDS are presented at **https://www.census.gov/programs-surveys/bds.html**.

The data shown in Figure INV-26 are created by summing the young firms across years zero to five and by summing employment created by these firms across years zero to five.

List

OSI-Approved and Machine Detectable Licenses

- SPDX: Software Package Data Exchange
- 0BSD: BSD Zero Clause License
- AFL-3.0: Academic Free License v3.0
- AGPL-3.0 GNU: Affero General Public License v3.0
- Apache-2.0: Apache License 2.0
- Artistic-2.0: Artistic License 2.0
- BSD-2-Clause: BSD 2-Clause "Simplified" License

- BSD-3-Clause: BSD 3-Clause "New" or "Revised" License

- BSL-1.0: Boost Software License 1.0

- CECILL-2.1: CeCILL Free Software License Agreement v2.1

- ECL-2.0 Educational Community License v2.0

- EPL-1.0: Eclipse Public License 1.0

- EPL-2.0: Eclipse Public License 2.0

- EUPL-1.1: European Union Public License 1.1

- EUPL-1.2: European Union Public License 1.2

- GPL-2.0: GNU General Public License v2.0 only

- GPL-3.0: GNU General Public License v3.0 only

- ISC: ISC License

- LGPL-2.1: GNU Lesser General Public License v2.1 only

- LGPL-3.0: GNU Lesser General Public License v3.0 only

- LPPL-1.3c: LaTeX Project Public License v1.3c

- MIT: MIT License

- MPL-2.0: Mozilla Public License 2.0

- MS-PL: Microsoft Public License

- MS-RL: Microsoft Reciprocal License

- NCSA: University of Illinois/NCSA Open-Source License

- OFL-1.1: SIL Open Font License 1.1

- OSL-3.0: Open Software License 3.0

- PostgreSQL: PostgreSQL License

- UPL-1.0: Universal Permissive License v1.0

- Unlicense: The Unlicense

- Zlib: zlib License

References

Aksoy AY, Beaudry C. 2021. How Are Companies Paying for University Research Licenses? Empirical Evidence from University-Firm Technology Transfer. *The Journal of Technology Transfer* 46:2051-121. Available at **https://doi.org/10.1007/s10961-020-09838-x**. Accessed October 2021.

Audretsch DB, Link AN, Scott JT. 2002. Public/Private Technology Partnerships: Evaluating SBIR-Supported Research. *Research Policy* 31(1)145–158.

AUTM. 2021. Statistics Access for Technology Transfer Database (version 4.2). Accessed February 2021.

Carlino GA, Chatterjee S, Hunt RM. 2006. *Urban Density and the Rate of Invention.* Working paper 04-16. Available at **https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2006/wp06-14.pdf?la=en&hash=7F943B98EDA4192D0F163B95116D4582**. Accessed October 2021.

Carlino GA, Chatterjee S, Hunt RM. 2007. Urban Density and the Rate of Invention. *Journal of Urban Economics* 61(3):389–419. Available at **https://doi.org/10.1016/j.jue.2006.08.003**. Accessed 10 June 2021.

de Rassenfosse, G, Dernis H, Guellec, D Piccic, L, and Pottelsberghe de la Potteriee, B. 2013. The Worldwide Count of Priority Patents: A New Indicator of Inventive Activity. Research Policy 42(3):720–737.

Dworak-Fisher K, Mirel L, Parker J, Popham J, Prell, M, Schmitt, R, Seastrom, M, Young L. 2020. A Framework for Data Quality. FCSM 20-04. Federal Committee on Statistical Methodology. Available at **https://nces.ed.gov/fcsm/pdf/FCSM. 20.04_A_Framework_for_Data_Quality.pdf**. Accessed September 2021.

Gousios G, Spinellis D. 2012. *GHTorrent: Github's Data from a Firehose.* 9th IEEE Working Conference on Mining Software Repositories (MSR). Available at **https://ieeexplore.ieee.org/document/6224294**. Accessed 9 June 2021.

Lerner J, Nanda R. 2020. *Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn.* NBER Working Paper 27492. Available at **http://www.nber.org/papers/w27492**. Accessed September 2021.

National Institute of Standards and Technology (NIST), U.S. Department of Commerce. 2019. *Federal Laboratory Technology Transfer Fiscal Year 2016: Summary Report to the President and the Congress.* Available at **https:// www.nist.gov/system/files/documents/2019/10/30/fy2016_fed_lab_tech_transfer_rept_fina_9-10-19.pdf**. Accessed 9 June 2021.

Organisation for Economic Co-operation and Development (OECD), Eurostat. 2005. *Oslo Manual 2005: Guidelines for Collecting, Reporting and Using Data on Innovation*. 3rd ed. Paris: OECD Publishing; Luxembourg: Eurostat.

Organisation for Economic Co-operation and Development (OECD), Eurostat. 2018. *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation*. 4th ed. Paris: OECD Publishing; Luxembourg: Eurostat.

Science-Metrix. 2022. *Patent and Trademark Indicators for the Science and Engineering Indicators 2022: Technical Documentation.* Available at **https://www.science-metrix.com/patent-and-trademark-indicators-for-the-science-and-engineering-indicators-2022-technical-documentation/**.

U.S. Department of Commerce, Census Bureau. 2021. New Vintage 2020 Population Estimates Available for Cities and Towns, and Housing Unit Estimates for the Nation, States and Counties. Available at **https://www.census.gov/newsroom/ press-releases/2021/2020-vintage-population-estimates.html**. Accessed July 2021.

U.S. Patent and Trademark Office (USPTO). 2015. *General Information Concerning Patents*. Available at **https:// www.uspto.gov/patents/basics**. Accessed 7 June 2021.

U.S. Patent and Trademark Office (USPTO). 2021. *Patenting in Technology Classes Breakout by Origin, U.S. Metropolitan and Micropolitan Areas: Explanation of Data*. Available at **https://www.uspto.gov/web/offices/ac/ido/oeip/taf/cls_cbsa/ explan_cls_cbsa.htm.** Accessed May 2021.

von Graevenitz G, Graham SJH, Myers AF. 2021. Distance (Still) Hampers Diffusion of Innovations. *Regional Studies.* Available at **https://www.tandfonline.com/doi/full/10.1080/00343404.2021.1918334**. Accessed October 2021.