Innovation

Invention, Knowledge Transfer, and Innovation

Technical Appendix

NSB-2020-4

January 15, 2020

Table of Contents

Technical Appendix

Patent Data Analytics and Terminology

The "Invention, Knowledge Transfer, and Innovation" report of *Science and Engineering Indicators 2020* uses the U.S. Patent and Trademark Office's (USPTO's) PatentsView database and the European Patent Office's (EPO's) Worldwide Patent Statistical Database (PATSTAT) as data sources for many of the analyses on patents and trademarks it presents. The National Center for Science and Engineering Statistics (NCSES), SRI International, and Science-Metrix prepared the tabulations from these databases presented in this report. A more detailed methodology report on the preparation of patent indicators by Science-Metrix is also available (**http://www.science-metrix.com/?q=en/publications/reports#/?q=en/publications/reports/**). The report also uses World Intellectual Property Organization (WIPO) data directly for three figures.

For the detailed tables by technology area in this report, patents are classified under WIPO's classification of 35 technical fields, based on International Patent Classification (IPC) codes. This report uses IPC reformed codes to prepare the patent data; these codes incorporate changes made with the eighth revision of the WIPO classification in 2006.

For the tabulations in this report, patents are assigned to geographic locations based on the country of residence of inventors. For patents awarded to inventors in multiple locations, this report uses fractional counts of patents to avoid double-counting. With fractional counts, a country receives partial credit for a patent in proportion to the number of named inventors who reside in that country divided by all named inventors. Patents can be tagged with multiple IPC codes and can fall under multiple WIPO technology areas. This report also uses fractional counts to assign patents to multiple sectors and technological fields.

USPTO Data and PatentsView Database

Inventors around the world obtain protection for inventions through national and regional jurisdictions. The 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 patents protect inventions in the U.S. market.

The USPTO, in collaboration with other federal agencies and academic institutions, maintains the data analysis and visualization platform PatentsView. 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 the WIPO technology fields. PatentsView also applies statistical techniques to match and link inventor names and locations. This matching of names and locations, known as *inventor and location disambiguation*, is important because it enables analyses of patterns and trends in patenting activity in the United States and abroad.

International Patent Families

The international patent family data presented in this report come from the EPO's PATSTAT (**https://www.epo.org/searching-for-patents/business/patstat.html#tab-1**). A *patent family* refers to a group of related patents that has one invention in common. Patent families provide a broad unduplicated measure of global invention.

International Patent Documentation patent families across all patent offices covered in PATSTAT are counted according to the year of the first granted patent in the patent family, also known as a priority patent. PATSTAT assigns patent families to geographic locations based on patent inventorship information found on the priority patent. PATSTAT allocates patent families among regions, countries, or economies using fractional counts based on the residences of all named inventors.

WIPO Patent Data

Data for WIPO-sourced figures are taken from the WIPO Statistics Database. This database is compiled primarily from WIPO's annual intellectual property (IP) statistics surveys (see below) and data compiled by WIPO in processing international applications and registrations through the Patent Cooperation Treaty (PCT), Madrid System, and Hague System. In cases where IP offices and countries do not provide statistics directly, WIPO uses data published on IP offices' websites or in their annual reports to supplement questionnaire responses. The application data include data on both direct filings and filings made through WIPO-administered international systems (where applicable). These data are available from WIPO's IP Statistics Data Center at **https://www.wipo.int/ipstats/en/**.

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) like 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, publication 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 in order to match NPL citations in PatentsView to cited publications appearing in Scopus.

Patent-Related Indicators in Indicators 2020

Patents Granted

This indicator reflects the number of patents granted to a country, sector, or organization. Patents have inventors (one or more) and grantees, in which the latter become the owners of the IP covered by the patent. This indicator primarily presents the fractional count of patents by inventor, although in some cases, it may present information by grantee; the notes for tables and figures specify the approach.

Patenting Activity Index

For any given technology field, the Patenting Activity Index indicates the extent to which a country specializes in that field. As a measure of specialization, or concentration of patenting activity, the Patenting Activity Index assesses the share of a country's patents produced in a given technological area relative to the global share of patents in that technological area. Back in 2016, for instance, the United States produced about 3,300 (or 2.3%) of its 143,000 patents in information technology (IT) methods for management (Table S8-4 and Table S8-11). By comparison, at the global level, only about 4,400 (or 1.4%) of 304,000 total patents were granted in IT methods for management (Table S8-4 and Table S8-11). Thus, the United States had a higher concentration of patenting activity in IT methods for management compared to the global average in 2016.

This indicator is indexed to 1.00; an index value above 1.00 shows that a country produces a larger share of its patent output in a given technological area compared to the global share of patent output in that area. Similarly, a value below 1.00 shows that a country produces a lower share of its patent output in a technological area compared to the global share of patent output in the same area. Whenever a country's share of patents in one area increases, its share in other areas must decrease proportionately.

The Patenting Activity Index does not include plant patents. Internationally comparable data on plant patents are not available. The USPTO grants plant patents to inventors who have invented or discovered and asexually reproduced a distinct and new variety of plant, other than a tuber-propagated plant or a plant found in an uncultivated state (USPTO 2015). In contrast, the EPO does not patent plants, animals, or biological processes for producing plants or animals (EPO 1998).

Trademark Data

Trademarks, which protect original symbols, are issued by national and regional offices. Trademark data used in this report come from three sources: WIPO, the USPTO, and the European Union Intellectual Property Office.

WIPO statistics cover annual trademark applications to the top five national offices by total number of applications.

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. For more information about this classification, see **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 in proportion to the number of trademark holders who reside in that country divided by all the trademark holders. This report also uses fractional counts to assign trademarks to the corresponding categories under the classification.

Plant Varieties

Plant variety data presented in this report come from the International Union for the Protection of New Varieties of Plants (UPOV)—an intergovernmental organization established by international convention to provide and coordinate plant variety protection. Plant variety protection covers nonhybridized plants, those created by seeds and tubers.

The plant variety protection data presented in this report include official statistics reported by 75 national offices to UPOV (UPOV 2019). These data are complete for 43 of these offices; 11 offices are missing data for a single year, 5 offices are missing data for 2 years, and the remaining 16 offices are missing data for more years. China includes Hong Kong and might also include Taiwan because it does not appear as a standalone country in the UPOV reports. Like patents and trademarks, this report allocates plant variety protections geographically and by category, according to holder information.

Intangible Investment Data

U.S. investment data in R&D; computer software; and entertainment, literary, and artistic originals reflect the accounting of intangible investment in the national income and product accounts (NIPA). Investment in artistic originals includes expenditures for the development and production of theatrical movies, long-lived television programs, books, music, and other artistic originals. This report shows these data in 2012 constant dollars, adjusting for the impact of price change. The Bureau of Labor Statistics (BLS) reports these data for use in its multifactor productivity measurement program. The Bureau of Economic Analysis (BEA) prepares the underlying investment data as part of its measurement of produced fixed assets.

PitchBook Venture Capital Data

The venture capital data used in this report supplement the analysis by shedding light on trends of emerging technologies and products. These proprietary data come from PitchBook Data, Inc., a company that collects financial and business data on the Web and provides subscription-based data (**https://pitchbook.com/**). Table S8-63 provides definitions of the PitchBook venture capital investment categories.

Business Innovation Survey Data

The Oslo Manual, prepared by the Organisation for Economic Co-operation and Development (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 (CIS) from the European Union (EU) Statistical Office and the Business R&D and Innovation Survey (BRDIS) from NCSES and the National Science Foundation. Following *The Oslo Manual*, these surveys define innovation as "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.

Community Innovation Surveys (CIS)

The CIS represent a coordinated effort to collect comparable innovation data across EU countries. They are conducted in the 28 EU member nations and used as the basis for other countries' data collection. The European Commission coordinates and integrates data collection for the EU member nations. OECD also uses these data in its international comparisons for the Science, Technology and Industry Scoreboard (**https://www.oecd.org/innovation/inno/inno-stats.htm#indicators**).

Business R&D and Innovation Survey (BRDIS)

BRDIS, described in the *Indicators 2020* report "Research and Development: U.S. Trends and International Comparisons," includes innovation questions derived from *The Oslo Manual* and the CIS. This survey has collected innovation data for nonfarm U.S. private industries with five employees or more since 2008.

OECD Innovation Data

The innovation rates across countries are extracted from OECD's Science, Technology and Industry Scoreboard: **https://www.oecd.org/innovation/inno/inno-stats.htm**.

Productivity Data

Growth in multifactor productivity and *total factor productivity* refer to the same concept: the growth in economic output not accounted for by changes in the quantity and quality of the inputs. In contrast, *labor productivity growth* is the growth in the amount of goods and services produced (output) compared with the growth in the hours worked to produce them. BLS calculates both for the United States.

The methodology for producing the U.S. productivity data used in this report is found at **https://www.bls.gov/mfp/**. BLS's sources for the productivity data are BLS data and NIPA data from BEA.

For international comparisons of productivity, this report uses data extracted from the OECD Productivity Database. OECD compiles and harmonizes these data by accounting for methodological changes in national accounts' statistics, such as the revision to the System of National Accounts 2008, and changes in international industrial classifications (**https://stats.oecd.org/**).

BLS calculates growth as the average annual rate of growth between the first year and the last year of each period.

Business Dynamics Statistics (BDS)

BDS is a public-use data set 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**.

References

European Patent Office. 1998. *Guidelines for Patent Examination*. **https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_ii_5_4_1.htm**. Accessed 3 November 2019.

International Union for the Protection of New Varieties of Plants (UPOV). 2019. *Plant Variety Protection Statistics*. **https://www.upov.int/meetings/en/topic.jsp?group_id=251**. Accessed 10 June 2019.

Organisation for Economic Co-operation and Development (OECD), Eurostat. 2005. *Oslo Manual: Guidelines for Collecting, Reporting and Using Innovation Data*. 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 Innovation Data*. 4th ed. Paris: OECD Publishing; Luxembourg: Eurostat.

U.S. Patent and Trademark Office (USPTO). 2015. *General Information Concerning Patents*. Available at **https://www.uspto.gov/patents-getting-started/general-information-concerning-patents#heading-31**. Accessed 3 November 2019.