


NATIONAL SCIENCE BOARD

SCIENCE & ENGINEERING INDICATORS 2020

Industry Activities

Production and Trade of Knowledge- and Technology-Intensive Industries

Technical Appendix

NSB-2020-5

January 23, 2020

This publication is part of the *Science and Engineering Indicators* suite of reports. *Indicators* is a congressionally mandated report on the state of the U.S. science and engineering enterprise. It is policy relevant and policy neutral. *Indicators* is prepared under the guidance of the National Science Board by the National Center for Science and Engineering Statistics, a federal statistical agency within the National Science Foundation. With the 2020 edition, *Indicators* is changing from a single report to a set of disaggregated and streamlined reports published on a rolling basis. Detailed data tables will continue to be available online.

Table of Contents

List of Tables

List of Figures

Technical Appendix

Industry Data, Methodology, and Terminology

The data and indicators presented in this report permit the tracing and analysis of broad patterns and trends that shed light on the spread and shifting distribution of global knowledge- and technology-intensive (KTI) capabilities. This report uses a variety of data sources including international industry production data from IHS Markit, trade data compiled by Oxford Economics, trade in value-added data from the Organisation for Economic Co-operation and Development (OECD), and a variety of other data sources on artificial intelligence (AI). This appendix provides information on the data sources utilized in this report. Internationally comparable data are generally derived from multiple national sources and are prone to varying issues of quality and reliability that necessitate caution when making international comparisons.

IHS Comparative Industry Service Forecast Database

The international industry production data are drawn from a proprietary database compiled by IHS Markit. The IHS Comparative Industry Service (CIS) Forecast Database provides consistent coverage for over 70 countries and over 100 industrial sectors for several macroeconomic indicators including the value-added output data presented in this report. The primary data sources on industry output are the National Income Accounts (NIA) from the countries' national statistical agencies and data from cross-national organizations including the Industrial Structure Statistics from the OECD Structural Analysis (STAN) database, the International Yearbook of Industrial Statistics from the United Nations Industrial Development Organization (UNIDO), and the National Accounts Statistics from the United Nations System of National Accounts (UNSNA).

IHS Markit compiles the data in the CIS database using a tiered approach where data from the OECD, UNSNA, and UNIDO form the foundation for all sectors and most countries. These are harmonized data and provide consistency and comparability across countries and across time. The OECD STAN database provides data for the OECD member countries. For countries not included in the OECD database, and for industries whose coverage in STAN is not sufficiently detailed, IHS Markit combines the OECD data with the UNSNA and UNIDO databases. The UNSNA contains national accounts data for most countries in the world and facilitates macroeconomic comparisons among national economies at the broad industry level. Data availability, however, varies across countries and fiscal years as not all United Nations (UN) member countries are able to provide a complete set of data. The UNIDO data set provides highly disaggregated data for the manufacturing sector. Both the UN and OECD data are further supplemented by other international sources such as the International Labor Organization (ILO) and Eurostat and by individual country sources.

For some countries or economies, IHS Markit collects more granular and timely data directly from their national statistical agencies. These countries or economies include the United States (Bureau of Economic Analysis), Brazil (Statistics Brazil), China (China National Bureau of Statistics) and Taiwan (Taiwan Directorate-General of Budget, Accounting & Statistics [DGBAS] and National Statistics). Finally, IHS Markit brings the collected data forward in time as needed using data from individual country sources, global trade associations, and other sources to provide timely measures of industry-level business activity.

Economic units within a company are classified according to the UN's International Standard Industrial Classification of All Economic Activities (ISIC Rev. 4). The criteria used to define and delineate the economic activities under this classification are based on similarities in inputs and factors of production, the process and technology of production, and characteristics and use of outputs. For economic units that engage in one type of activity, classification of the unit is straightforward. For economic units that engage in several types of independent activities, considerable proportions of the activities of the unit may be included in more than one class of ISIC. If the unit cannot be split into separate statistical units based on these activities, it is classified based on the principal activity, the activity that contributes most to the value added of the unit. Thus, a company that primarily manufactures pharmaceuticals, for example, but also operates a retail business may have all its economic activity counted under pharmaceuticals.[1]

Finally, the value-added output data discussed in this report are presented in nominal values; the value of production is measured at current market prices and it is not adjusted for inflation. In addition, value-added output is measured in current dollars. For countries outside the United States, value added is recorded in the local currency and converted at the prevailing nominal market exchange rate. This choice comes with some limitations, particularly when an economy's currency exchange rates are not market determined.

Oxford Economics Global Trade Database

The Oxford Economics Global Trade database contains detailed information on bilateral trade flows for over 30 major economies including individual product flows for almost 70 product categories. The historical data on trade in goods disaggregated by product are ingested directly from the UN Comtrade database. This database is the world's largest database of trade statistics and is maintained by United Nations Statistics Division (UNSD). It provides information on annual trade flows in nominal U.S. dollar terms, with products classified according to the Standard International Trade Classification (SITC) system.

Oxford Economics cross-validates the trade data with other sources including the International Trade Center (ITC) Trade Map to address missing observations and other discrepancies. The ITC Trade Map is based primarily on UN Comtrade, but it is supplemented with national and regional sources. The ITC Trade Map also contains trade statistics for countries that do not report their trade data to UN Comtrade; these statistics are generated with data reported by partner countries. In addition to data cleanup and cross-validation, Oxford Economics also estimates the latest year of the data because the UN Comtrade database reports trade flows with a lag; annual data for many countries are not available for several months after the end of the year. These estimates use available monthly data along with trends in exports/import flows, intermediate and final demand, and production by sector.

Trade in Value Added Database

Trade data in value-added terms provide a comprehensive picture of a geographic location's economic contribution to trade flows and this report uses a database created for this purpose by OECD and the World Trade Organization (WTO). Goods and services are increasingly produced in global value chains (GVCs). A GVC consists of the full range of activities that take place across the globe to bring a product or service from its conception to its final form for end use by consumers. Countries participate in global production through GVCs by specializing in segments of production in which they have a comparative advantage. Value-added trade data allow this contribution to be measured separately from the total value of the traded products.

OECD started publishing trade in value added (TiVA) statistics in 2013 through a joint initiative with WTO. The OECD TiVA statistics rely on a global input-output database, a data set that links production processes within and across countries. The global input-output database is derived by combining harmonized single-country input-output tables with bilateral trade data. The input-output tables are sourced from national statistical agencies and harmonized by OECD. They show production of output by all sectors and the allocation of domestic output among intermediate and final uses, including exports. The linkage of input-output data with trade data captures the bilateral exchanges of intermediate goods and services.

For countries in the TiVA database, the overall trade balance is consistent with official national accounts figures. However, bilateral trade figures may differ from those published by national statistical agencies; the latter result in inconsistent global trade figures because of differences across countries in treatment of various aspects of trade such as re-exports and transit trade. The international statistics community continues work to improve consistency in measuring international trade flows, particularly in services, where there are substantial differences across national statistics.[2]

The most recent update (2018) of the TiVA database covers 64 economies, including the OECD countries, the European Union (EU) and the G20 countries, and most East and Southeast Asian economies for the years 2005–15 with projections for 2016.[3] The input-output tables from which the TiVA indicators are derived are based on the 2008 System of National Accounts (SNA 2008). Indicators are available for 36 industries within a hierarchy based on ISIC Rev. 4.

Along with its advantages, trade in value added presents some measurement challenges. The value added of companies with diversified businesses may be assigned to the single industry that accounts for the largest share of the company's business. A company classified as manufacturing may include services, and vice versa. Disentangling the domestic and foreign content in GVCs is further complicated by fragmentation of production within multinational enterprises and trade in inputs that are further traded as more processed inputs.

Classification of Industries Based on Research and Development Intensity

This *Indicators 2020* report relies on a classification of industries based on research and development (R&D) intensity developed by Galindo-Rueda and Verger (2016). This classification represents both an update and an extension to prior OECD taxonomies; it incorporates ISIC Rev. 4 and extends for the first time the analysis to non-manufacturing industries including a broad range of services.

The OECD study by Galindo-Rueda and Verger (2016) classifies industries into R&D intensity groups based on a measure of R&D performance intensity computed as the ratio of each industry's business R&D expenditures to the industry's gross value added. Value added is a net measure of output; it is the difference between the total revenue generated from the sale of an industry's output and the total cost of inputs that were used in production such as the cost of labor, raw materials, and services purchased from other businesses. Value added is used in the computation of R&D intensity instead of gross output because it avoids double counting of intermediate production.

This R&D intensity measure only captures the R&D that is directly performed by industries. It does not capture any R&D that is indirectly incorporated in an industry's purchases of intermediate inputs and capital goods. For any given industry, the R&D intensity is computed as a weighted average of corresponding industry R&D intensities for a core sample of countries using value-added output in purchasing power parities (PPP) as weights.[4] R&D intensities are computed using data from 2011. Because this taxonomy is based on average R&D intensities across countries, the R&D intensities for industries in individual countries are likely to vary from the average.

Industries are classified by OECD into five R&D intensity groups: high, medium-high, medium, medium-low, and low. For KTI industries in the high and medium-high R&D intensity groups, the R&D expenditures as a share of value added range from almost 6% (railroad, military vehicles and transport not elsewhere classified [nec]) to over 30% (air, spacecraft and related machinery). Among these are three service industries: scientific R&D services (30.4%), software publishing (28.9%), and IT and other information services (5.9%). The full classification of industries is presented in **Table SA6-1**.

TABLE SA6-1

OECD classification of industries by R&D intensity

(List of industries)

R&D Intensity Category	Manufacturing			Services		
	Industry code	Name	R&D intensity (%)	Industry code	Name	R&D intensity (%)
High R&D intensity	303	Air and spacecraft and related machinery	31.69	72	Scientific research and development	30.39
	21	Pharmaceuticals	27.98	582	Software publishing	28.94
	26	Computer, electronic, and optical products	24.05			

TABLE SA6-1

OECD classification of industries by R&D intensity

(List of industries)

R&D Intensity Category	Manufacturing			Services		
	Industry code	Name	R&D intensity (%)	Industry code	Name	R&D intensity (%)
Medium-high R&D intensity	252	Weapons and ammunition	18.87	62-63	IT and other information services	5.92
	29	Motor vehicles, trailers, and semi-trailers	15.36			
	325	Medical and dental instruments	9.29			
	28	Machinery and equipment nec	7.89			
	20	Chemicals and chemical products	6.52			
	27	Electrical equipment	6.22			
	30X	Railroad, military vehicles and transport nec (ISIC 302, 304, and 309)	5.72			
Medium R&D intensity	22	Rubber and plastic products	3.58			
	301	Building of ships and boats	2.99			
	32X	Other manufacturing except medical and dental instruments	2.85			
	23	Other non-metallic mineral products	2.24			
	24	Basic metals	2.07			
	33	Repair and installation of machinery and equipment	1.93			
Medium-low R&D intensity	13	Textiles	1.73	69-75X	Professional, scientific and technical activities except scientific R&D services (ISIC 69-75 less 72)	1.76
	15	Leather and related products	1.65	61	Telecommunications	1.45
	17	Paper and paper products	1.58	05-09	Mining and quarrying	0.8
	10-12	Food products, beverages and tobacco	1.44	581	Publishing of books and periodicals	0.57
	14	Wearing apparel	1.40			
	25X	Fabricated metal products except weapons and ammunition (ISIC 25 less 252)	1.19			
	19	Coke and refined petroleum products	1.17			
	31	Furniture	1.17			
	16	Wood and products of wood and cork	0.70			
	18	Printing and reproductions of recorded media	0.67			
Low R&D intensity				64-66	Financial and insurance activities	0.38
				35-39	Electricity, gas, and water supply, waste management, and remediation	0.35
				59-60	Audiovisual and broadcasting activities	0.32
				45-47	Wholesale and retail trade	0.28
				01-03	Agriculture, forestry and fishing	0.27
				41-43	Construction	0.21
				77-82	Administrative and support services	0.18
				90-99	Arts, entertainment, repair of household goods, and other services	0.11
				49-53	Transportation and storage	0.08
				55-56	Accommodation and food service activities	0.02

TABLE SA6-1

OECD classification of industries by R&D intensity

(List of industries)

R&D Intensity Category	Manufacturing			Services		
	Industry code	Name	R&D intensity (%)	Industry code	Name	R&D intensity (%)
				68	Real estate activities	0.01

nec = not elsewhere classified.

OECD = Organisation for Economic Co-operation and Development.

Note(s):

R&D intensity are global R&D expenditures as a share of global value added output of industry. The global R&D and value added excludes several economies, including China, India, and Mexico due to incomplete missing industry and R&D data. Industries classified according to International Standard Industry Classification, Revision 4.

Source(s):

Galindo-Rueda F, Verger F. 2016. OECD Taxonomy of Economic Activities Based on R&D Intensity. OECD Science, Technology and Industry Working Papers, 2016/04, OECD Publishing, Paris.

This classification of industries is sensitive to the choice of the group of countries. The OECD analysis underlying this classification does not fully capture global production and R&D because of the exclusion of several economies that have incomplete data, including Brazil, India, and China. Because these missing economies may have large global production shares and R&D intensities that may be substantially different from the average of the core sample of economies included in the OECD analysis, their exclusion may result in differences in classification of R&D intensity.

Including partial data from China, for example, decreases the average R&D intensity in more than two-thirds of the industries. This is consistent with China producing a large share of global output for these industries, while investing a below sample average share in R&D. This is particularly the case for computer, electronic, and optical products; transport equipment; and chemicals and pharmaceuticals. In terms of impact on classification, the inclusion of China, for example, would downgrade the computer, electronic, and optical products classification from the high R&D intensive group to the medium-high R&D intensive group. The impact on R&D intensities and the classification of other industries is presented by Galindo-Rueda and Verger (2016:25).

The R&D intensities are also sensitive to the national accounting conventions used to compile and report the output data. The output data are compiled and reported under the 1993 SNA, prior to the capitalization of R&D expenditures with the SNA 2008. The treatment of R&D expenditures as an investment directly increases the measure of value added and decreases the R&D intensities; the more intensive the industries, the higher the downward impact in the R&D intensities. In a series of robustness checks, the authors estimate that with the SNA 2008 standard, the R&D intensities drop by 20% in the highest group, 6% in the medium-high group, 3% in the medium and medium-low groups, and 1% in the low group. The classification of industries in these groups, however, remains stable.[5]

Measurement of Skilled Technical Workforce in KTI Industries

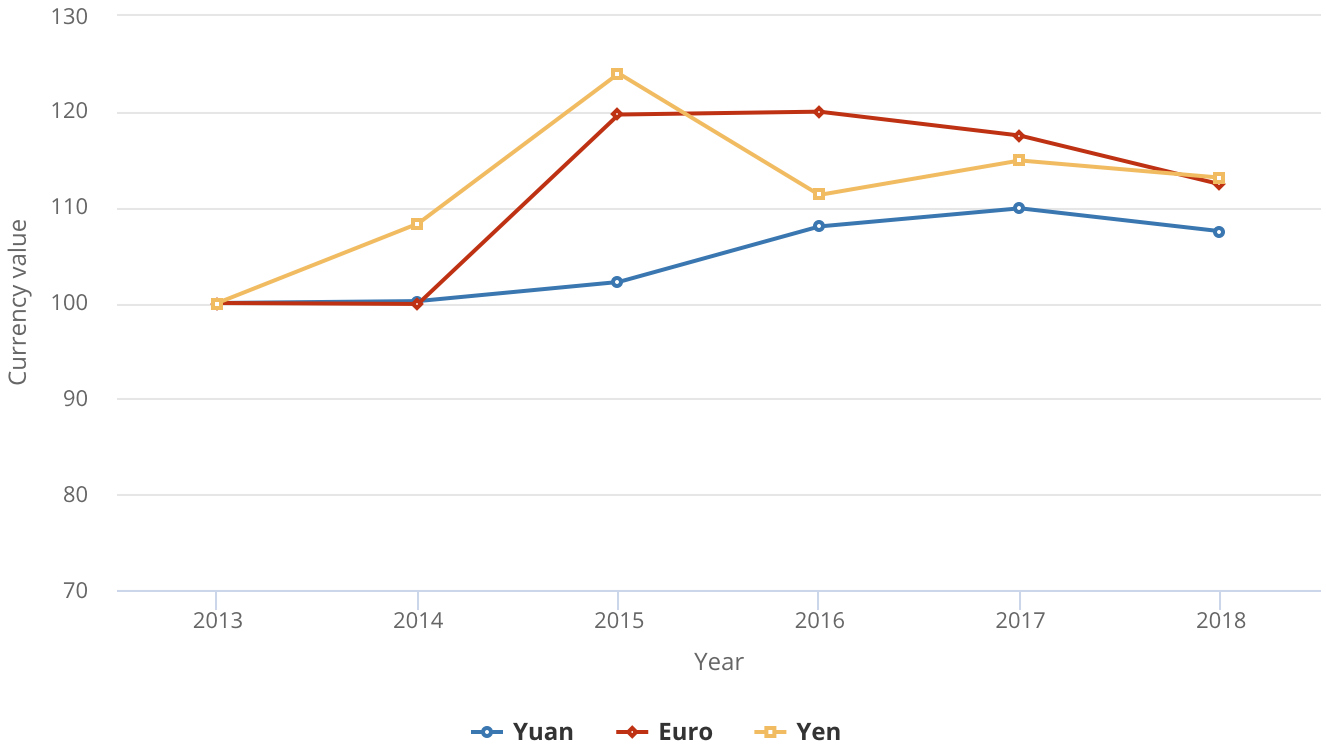
KTI industries are mapped to North American Industrial Classification System (NAICS) 2012 codes, and in turn to American Community Survey (ACS) industry codes, which are based on the NAICS. The mapping used U.S. Census Bureau's Industry and Occupation Code Lists & Crosswalks to produce the industry-specific skilled technical workforce (STW)-related statistics in this report (U.S. Census Bureau 2019). Information on occupations included in the STW and the methodology can be found in *Indicators 2018* report "Scientific and Engineering Labor Force."

Currency Exchange Rates of Major Economies

Between 2013 and 2018, the exchange rates of the world's four largest economies—China, the EU member countries that use the euro (the eurozone), Japan, and the United States—exhibited some fluctuations (**Figure SA6-1**) with depreciation of the yuan (8%), euro (12%), and Japanese yen (13%) against the dollar. The depreciation reduced the value of the KTI industries and other economic activities in dollar terms relative to their local currency in China, the eurozone, and Japan, resulting in higher growth rates in local terms relative to dollar terms (**Figure SA6-2**).

FIGURE SA6-1

U.S. dollar exchange rate with selected currencies: 2013–18



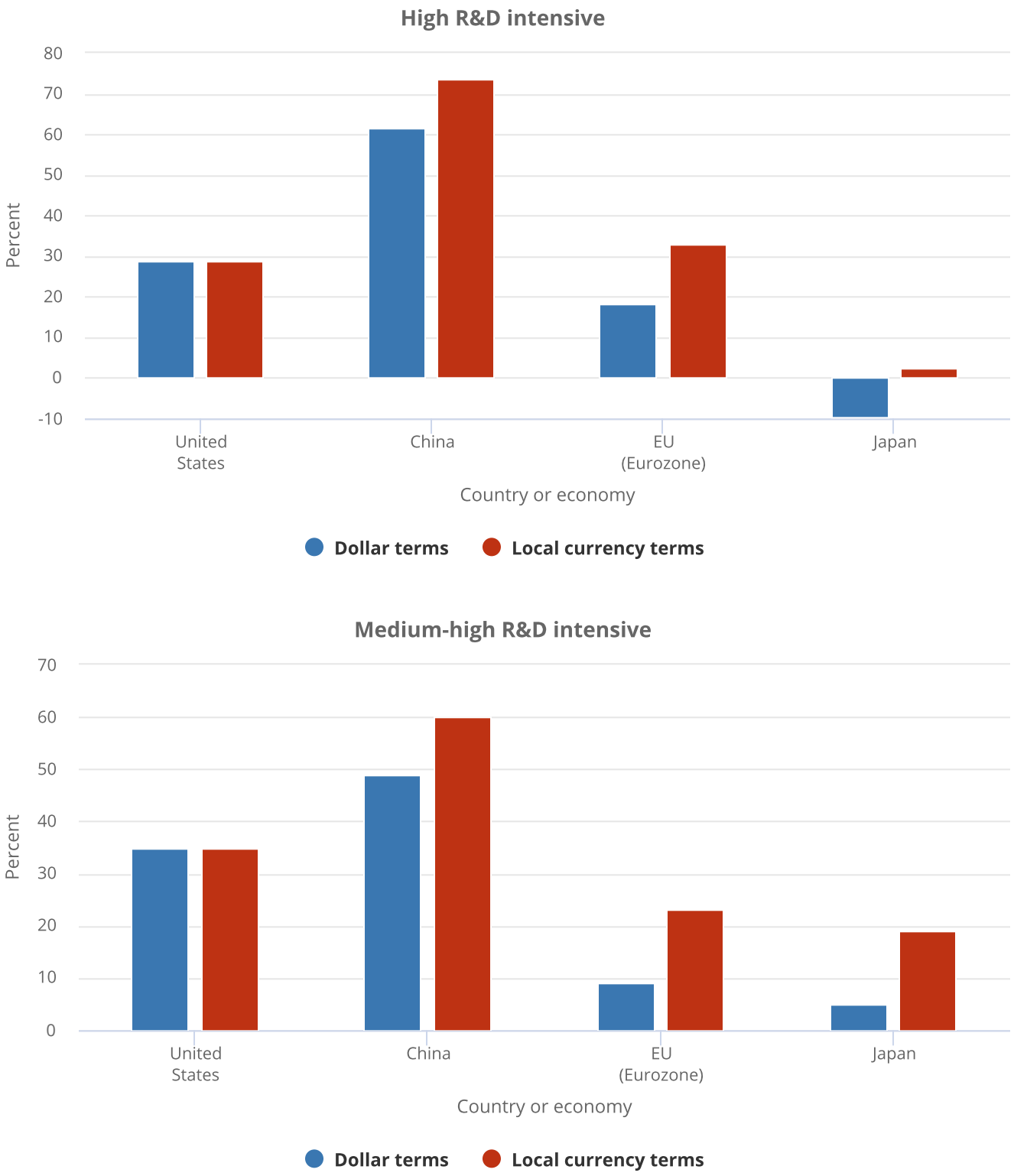
Note(s):
Currency value is expressed as an index of 100 in 2013.

Source(s):
Federal Reserve, Economic and Research and Data, Foreign Exchange Rates.

Science and Engineering Indicators

FIGURE SA6-2

Growth in output of selected categories of industries, by selected country or economy: 2013–18



EU = European Union.

Note(s):
Value added is the amount contributed by a country, firm, or other entity to the value of a good or service and excludes purchases of domestic and imported materials and inputs. High and medium-high R&D industries are classified by the Organisation for Economic Co-operation and Development. High R&D intensive industries include aircraft; computer, electronic, and optical products; pharmaceuticals; scientific research and development; and software publishing. Medium-high R&D intensive industries include motor vehicles and parts; medical instruments, electrical equipment; machinery and equipment; chemicals excluding pharmaceuticals: and railroad, military vehicles, and other transport. EU member countries comprising the Eurozone consist of Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Slovakia, Slovenia, Spain, and Sweden. See Table S6-3 and Table S6-4.

Source(s):
Federal Reserve, Economic Research and Data, Foreign Exchange Rates; IHS Markit, special tabulations (2019) of Comparative Industry Service.

International comparisons of industry, trade, investment, and other global economic activities often use current dollars at market exchange rates. Most global economic activities are dollar denominated, which facilitates comparison. In addition, many economists believe that market exchange rates reflect, at least to some degree, differences in economic performance among various countries (Balke, Ma, and Wohar 2013:2).

However, fluctuations in exchange rates may also reflect factors other than economic performance. Governments influence the level of their exchange rate indirectly through macroeconomic policies and directly through buying and selling currencies. In addition, factors such as political instability or the short-term effects of global financial events on a country's economy can cause currency fluctuations that are unrelated to differences in national economic performance. Comparing economic data from different countries in current dollar terms provides a broadly indicative reflection of a country's relative economic performance.

Artificial Intelligence Indicators

The AI Index 2018 is the data source for several of the indicators in this report; it is compiled by the Human-Centered AI Initiative at Stanford University (Shoham et al. 2018). In this appendix we describe the data sources used in this report; however, more complete methodologies are provided with the AI Index Report at **https://aiindex.org/**.

AI-Related Scientific Publications

AI-related publication data in the report include the quantity of AI-related publications produced each year, the relative citation impact of publications from different geographic locations, and the areas of concentration of AI-related publications. These data come to the AI Index Report from a 2018 Elsevier report on analysis of data in the Scopus database (Shoham et al. 2018; Elsevier 2018). This is a separate analysis from that described in the *Indicators 2020* report "Publications Output: U.S. Trends and International Comparisons," though the underlying database is the same.

For the AI-related work described in this report, Elsevier used a static version of the database across the years 1996–2017 and OECD fields of R&D are used to classify the publication fields in that report and focus on domains where AI is applied. The Elsevier report identified the following categories: machine learning and probabilistic reasoning; neural networks; computer vision; search and optimization; natural language processing and knowledge representation; fuzzy systems; and planning and decision making (Elsevier 2018).

Adoption of AI Capabilities in Industry

Data on industry adoption of AI capabilities in this report also come from the 2018 AI Index (Shoham et al. 2018). This information is based on a McKinsey and Company survey of 2,135 respondents across the world. To estimate the country-level statistics, the McKinsey and Company study weighted country responses to control for GDP (McKinsey and Company 2017, 2017b). Shoham et al. (2018) provides more information on the survey.

AI Skills in Job Openings

Data on job openings for AI-related skills were compiled by Monster.com based on resume data that was analyzed by Garner TalentNeuron, a private sector data and analysis firm (Shoham et al. 2018), and are not weighted for the size of the U.S. workforce. To identify skills related to AI, the search keyword "artificial intelligence" was paired with the following topics to obtain the results: natural language processing, machine learning techniques, deep learning, computer vision, and speech.[6]

PwC 2018 Global Innovation Study

R&D spending related to AI are drawn from the *2018 Global Innovation 1000 Study,* produced by Strategy& (**https://www.strategyand.pwc.com/innovation1000**), a PwC company. Strategy& collects publicly available data on company R&D spending, and adjusts it to remove double counting for subsidiary firms. These data are supplemented for the top 1,000 companies with additional data from Capital IQ and Thomas Eiken and, according to the methodology, adjusted for internal and cross-national consistency.

PitchBook Venture Capital Data

The venture capital data shown in Table 6-4 are from PitchBook Data, Inc, a private sector financial services company that collects financial and business data on the Web and provides subscription-based data (**https://pitchbook.com/**). The categories of PitchBook investment are listed in Table S8-63, where the data are used more extensively than in this report.

Artificial intelligence is defined for PitchBook investment reporting as companies developing technologies that enable computers to autonomously learn, deduce and act, through utilization of large data sets. The technology enables development of systems that collect and store massive amounts of data and analyze that content to make decisions based on probability and statistical analysis. Applications for Artificial Intelligence & Machine Learning include speech recognition, computer vision, robotic control, and accelerating processes in the empirical sciences where large data sets are essential, such as gene-sequencing in life sciences (PitchBook and NVCA 2019).

References

Balke NS, Ma J, Wohar ME. 2013. The Contribution of Economic Fundamentals to Movements in Exchange Rates. *Journal of International Economics* 90(1):1–16. Available at **http://www.sciencedirect.com/science/article/pii/S0022199612001675**. Accessed 29 July 2019.

Galindo-Rueda F, Verger F. 2016. OECD Taxonomy of Economic Activities Based on R&D Intensity. *OECD Science, Technology and Industry Working Papers*, No. 2016/04, OECD Publishing, Paris, **https://doi.org/10.1787/5jlv73sqqp8r-en**. Accessed 18 April 2019.

Elsevier. 2018. *Artificial Intelligence: How Knowledge is Created, Transferred, and Used.* **https://www.elsevier.com/research-intelligence/resource-library/ty/ai-report**. Accessed 31 July 2019.

McKinsey and Company. 2017. *Artificial Intelligence: The Next Digital Frontier?* **https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx**. Accessed 31 July 2019.

PitchBook Data, Inc., National Venture Capital Association (NVCA). 2019. *Venture Monitor 4Q 2018*. Seattle, WA: PitchBook Data, Inc. **https://pitchbook.com/news/reports/4q-2018-pitchbook-nvca-venture-monitor**. Accessed 14 March 2019.

Shoham Y, Perrault R, Brynjolfsson E, Clark J, Manyika J, Niebles JC, Lyons T, Etchemendy J, Grosz B, Bauer Z. 2018. *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University: Stanford, CA.

U.S. Census Bureau. 2019. Industry and Occupation Code Lists & Crosswalks. **https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html**. Accessed 15 November 2019.

Notes

1 More information on ISIC Rev. 4 is available at **https://unstats.un.org/unsd/publication/seriesm/seriesm_4rev4e.pdf**.

2 For a more detailed description of concepts and methodologies related to the TiVA statistics and a discussion of ongoing projects and initiatives, see OECD-WTO Joint Note. Trade in Value-added: Concepts, Methodologies, and Challenges (**https://www.oecd.org/sti/ind/49894138.pdf**, accessed 27 June 2019).

3 The G20 is an international group of 20 major economies. It consists of 19 individual countries and the European Union. The countries are Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Mexico, Russia, Saudi Arabia, South Africa, South Korea, Turkey, United Kingdom, and the United States.

4 The countries are 27 OECD countries and 2 partner countries or economies (Singapore and Taiwan). The OECD countries are: Australia, Austria, Belgium, Canada, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, South Korea, Mexico, the Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, the United Kingdom, and the United States.

5 See Galindo-Rueda and Verger (2016) for a full discussion of robustness checks.

6 The underlying data can be found at **https://docs.google.com/spreadsheets/d/1PENuyM-8RwjxaEOnfbY4I2mGZk_SZMffiUwtHCvMvcY/edit#gid=758065764**.