

Mapping high R&D city-regions worldwide A patent heat map approach

Stek, Pieter E.

10.1007/s11135-019-00874-w

Publication date

Document Version Final published version

Published in Quality and Quantity

Citation (APA)Stek, P. E. (2019). Mapping high R&D city-regions worldwide: A patent heat map approach. *Quality and Quantity*, *54*(1), 279-296. https://doi.org/10.1007/s11135-019-00874-w

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Mapping high R&D city-regions worldwide: a patent heat map approach

Pieter E. Stek¹

© The Author(s) 2019

Abstract

A methodology for identifying high R&D city-regions worldwide using patent data is presented. A heat map (kernel density estimation) approach is used which allows city-regions to be identified in areas with a high patenting intensity, a proxy for high R&D expenditure. The methodology is independent of any pre-existing administrative boundaries and it can therefore be applied to identify sub-national concentrations of R&D expenditure worldwide. This is an important advantage compared to other statistical data which is often only available at the national level. The results provide insight into the changing worldwide spatial distribution of R&D expenditure between 1997 and 2011, including the rapid rise of Asian R&D city-regions as well as less dramatic shifts among European and North American city-regions. The results also highlight some challenges of identifying high R&D city-regions and estimating R&D expenditure using patent data, and the existence of very large high R&D city-regions which encompass multiple cities. Some suggestions for improvement and further research are also proposed.

 $\textbf{Keywords} \ \ Innovation \cdot R\&D \cdot Cities \cdot Patents \cdot Mapping \cdot World \cdot Evolutionary \ economic \ geography$

1 Introduction

Like many other kinds of economic activity, Research and Development (R&D) is spatially concentrated but also globally distributed (Dicken 2007; Fujita et al. 2001; Malecki 2014). As a result most of the world's R&D expenditure is found in a relatively small number of city-regions around the world. Some of these city-regions are well known and have been extensively described in the scientific literature, such as California's Silicon Valley (Castells 2014; Saxenian 1996). But the global position of many other city regions is less clear. For example, while Seoul is well known as a global R&D hub, is the R&D activity of other cities in South Korea also globally significant? (Yoon and Park 2017) And what

Published online: 11 April 2019

Faculty of Technology, Policy and Management, Delft University of Technology, Gebouw 31, Jaffalaan 5, 2628 BX Delft, The Netherlands



Pieter E. Stek p.e.stek@tudelft.nl

about second-tier cities in Europe, the United States and Japan? Or emerging R&D centres in India and China? (Crescenzi and Rodríguez-Pose 2017).

A number of global city rankings have been produced in recent years, including the *Global Cities Report* by consulting firm A.T. Kearney, the *JLL Typology of World Cities* by the real estate firm Jones Lang Lasalle and the *The World According to GaWC* by Loughborough University's Globalization and World Cities Research Network (Beaverstock et al. 1999), but none of these focus exclusively on R&D or innovation activity. On the other hand there are numerous international rankings of innovation, including the *Global Innovation Index* (Dutta and Lanvin 2016), *Global Competitiveness Report* (Schwab and Sala-i-Martin 2015) and the *Bloomberg Innovation Index* by Bloomberg, a media company. These indexes all provide insight at the national level but ignore the sub-national city level. Therefore insight into the global innovation position of cities is lacking.

Because city-regions are the spatial scale at which innovation activity is concentrated, the purpose of this paper is to identify the world's most important city-regions in terms of R&D expenditure. Critical in this process is the use of patent data which is used (a) to identify high R&D city-regions, (b) as a proxy to estimate city-region R&D expenditure and (c) it allows the study of high R&D city-regions over longer timescales. These points are elaborated upon in the Methodology.

The ability to study city-regions over longer timescales allows global shifts in R&D expenditure to be observed (Dicken 2007) and it addresses the important theoretical question about the extent to which R&D activity is path-dependent or whether significant path-changing forces also exist (Boschma and Frenken 2006; Crescenzi and Rodríguez-Pose 2011). Three five-year periods from 1997 to 2011 are considered in the study and they show significant changes in the global hierarchy of innovation city-regions, especially the rise of city-regions in Asia. A table of the 20 city-regions with the greatest R&D expenditure is provided in the Results. Before that the theoretical background of the geography of R&D activity and global shifts is briefly discussed under Theory.

2 Literature review

R&D activity is highly globalised and spatially concentrated at the same time (Feldman 1994; Malecki 2014; Storper 1997), two features which at first may appear contradictory because globalisation suggests footlooseness, while spatial concentration suggests that specific territorial advantages do exist (Bathelt et al. 2004; Binz and Truffer 2017; Gertler and Wolfe 2006). Spatial concentration is driven by economies of scale effects, the lowering of transaction and collaboration costs and a greater variety of local suppliers, customers and skilled labour, as well as less tangible social-environmental factors such as trust, local competition and creativity (Capello 2009; Feldman and Kogler 2010; Porter 1998). However spatial concentration can also lead to negative economies of scale effects such as congestion, higher cost and excessive competition for customers and resources (Martin and Sunley 2003). The effect of geographic concentration on firm growth therefore appears to be ambiguous (Frenken et al. 2015; Grillitsch and Nilsson 2017), although spatial concentration does appear encourage firm entry (Frenken et al. 2015).

While there is ambiguity over the influence of spatial concentration on R&D activity, there appears to be strong path dependence, whereby spatial concentrations of R&D activity persist over long periods of time (Boschma and Frenken 2006; Crescenzi and Jaax 2017; Crescenzi and Rodríguez-Pose 2011). Path-creating factors, which could lead to



significant shifts in the spatial concentration of R&D activity are public policy intervention and foreign direct investment in R&D (Crescenzi and Jaax 2017; Dicken 2007). Investment in higher education and basic research, as well as targeted public–private sector R&D collaboration, play an important role in establishing and attracting high technology industries to a particular location (Lee and Lim 2001; Lee et al. 2009). In addition the need to distribute R&D activity globally is primarily driven by the expanding complexity of technology. Technological complexity requires increasingly specialised knowledge which may only be available in a few locations and competitive pressures which encourage firms to seek the best researchers and collaboration partners globally (Audretsch et al. 2014; Locke and Wellhausen 2014). The globalisation of R&D has been facilitated by improvements in communications technology and a fall in air transportation costs enabling frequent online and face-to-face contact (Gertler 2003; Maskell et al. 2006).

The speed at which the spatial concentration of R&D shifts and high R&D city-regions grow or decline, provides an indication of the degree of path-dependence and path-creation in the current global R&D system. Significant spatial changes would suggest that the current global research and innovation environment is in flux with significant path-creation occurring due to public intervention, international private sector investment or other factors.

3 Data and methodology

The methodology consists of two main components: the identification of innovation city regions and the estimation of R&D expenditure using on patent data. The two components are sequential in the sense that the identified innovation regions and their patent counts are used as proxies for the R&D expenditure estimation.

3.1 Identification of high R&D city regions

The identification of high R&D city regions is based on patent data obtained from the PatentsView database which is published by the Office of Chief Economist in the United States Patent and Trademark Office (USPTO) and contains data on 6,647,699 patent grants from the USPTO (May 2018 edition). Because the United States is a large economy, many foreign entitites also apply for patent protection at the USPTO, providing global coverage of patents, although with a home bias in favour of domestic United States patents. This home bias requires corrections to be made when comparing the US to other countries. The advantage of using a single source of patents (USPTO) is that all patents are granted in accordance to a single standard, improving the validity of making international comparisons (Toivanen and Suominen 2015).

Alternative databases to the USPTO database would be the European Patent Office (EPO) and the Japan Patent Office (JPO) as these also serve large advanced economies. However of the three databases, the USPTO database offers the greatest international coverage (Kim and Lee 2015), making it the most suitable for international comparative studies such as this one. In future China's State Intellectual Property Office (SIPO)

¹ The PatentsView database tables can be downloaded at: http://www.patentsview.org/download/ (accessed 24 March 2019).



database may be an alternative to the USPTO, as China's economic and technological importance continues to grow.

The USPTO database contains basic bibliographic information of patent documents such as patent identification numbers, application dates, inventors and assignees, the addresses of inventors and assignees, and patent citations, along with technological classifications. In this study the patent inventor addresses and application dates are used to identify the time and place where the R&D activity that resulted in the patent took place. The address allows for the identification of a geographic location of the inventive activity that led to the patent application.

The choice of inventors (individuals who carried out the R&D) rather than assignees (typically firms that financed the R&D) is not trivial. Inventors' location provides information about where the R&D took place whereas the assignee location provides information about who owns the inventions. Given the globalization of R&D activity, assignees and inventors are frequently found in different countries.

Patent data allows for the identification of high R&D city-regions worldwide based on the actual local patenting intensity (Alcácer and Zhao 2016; Duranton and Overman 2005). This is an advantage compared to other statistical data, which is often based on administrative boundaries. Examples of sub-national administrative boundaries are states (Germany), regions (Italy), provinces (Spain), departments (France), counties (England) and prefectures (Japan). In some cases specific statistical boundaries exist, such as the metropolitan statistical area (United States). The first disadvantage of using administrative boundaries is their fit compared to the phenomenon being studied. Administrative boundaries may not fit well with actual innovation activity, which could spill across boundaries and/or take place in only a small part of the administrative area. An aggregate statistic for the administrative area could therefore give an inaccurate picture of the real spatial concentration of innovation activity. A second disadvantage of administrative boundaries is the difference in scales between countries. For example, China's Guangdong province (179,800 km², 108.5 million people) cannot really be compared to the Netherlands' Zuid-Holland province (2818 km², 3.6 million people), even both are the most populous provinces in their respective countries. Such differences make global comparisons at the sub-national level very challenging.

High R&D city-regions can be defined as geographic concentrations of patenting activity. The identification of concentrations is based on the inventor address information obtained from the patent database. This address information is converted into coordinates through geocoding. For example, the address 'Delft, The Netherlands' is converted into the coordinates 51.9995142, 4.2938295.

Although the PatentsView database does provide coordinates for patent addresses, upon closer examination a number of these appear to be inaccurate because the coordinates are located in a different country than the address or the coordinates are only located at the country or state level, and not at that of a town or city, which is a problem in larger countries. For this reason approximately 6.5% of addresses are geocoded again, a process carried out in three stages.

In stage one all addresses in countries or territories which are less than 20,000 km² in size are automatically assigned coordinates. The largest among this group is New Caledonia (18,575 km²), also included are countries and territories such as Kuwait, Montenegro, Qatar, Cyprus, Puerto Rico, Luxembourg, Hong Kong and Singapore.



In stage two the open-source TwoFishes geocoding application is used (using index files updated on 2015-03-05).² TwoFishes is used and maintained by FourSquare Labs Inc., a company that operates a popular local search-and-discovery service mobile application. An important advantage of TwoFishes is that it is open source and therefore its search results are reproducable. Twofishes has been used in published and peer-reviewed scientific papers (Hamstead et al. 2018; Sessions et al. 2016) and it is listed in *The SAGE Handbook of Social Media Research Methods* (Sloan and Quan-Haase 2017).

In stage three the Google Maps geocoding application is used, which appears to have a higher success rate than TwoFishes, but it is a closed-source commercial service, making it potentially more difficult to reproduce results. Google Maps is also used in other scientometric studies (Leydesdorff and Persson 2010; Leydesdorff et al. 2014) and is listed in the aformentioned *Handbook* (Sloan and Quan-Haase 2017). The combination of geocoding techniques raises the number of addresses that can be accurately located from 93 to 96%.

After geolocating inventors, each identified location receives a weighting (PTW_i) based on the number of inventors with an address in that location (INV_{ij}) divided by the number of inventors of the patent $(INVT_i)$ and then summed for all patents (j). Thus:

$$PTW_i = \sum INV_{ij}/INVT_j$$

An example of the calculation: a patent with 3 inventors, 2 of whom have an address in 'Delft, The Netherlands' would therefore add a weighting of 2/3 = 0.67 to the location of 'Delft, Netherlands' (51.9995142, 4.2938295).

Once the geographic concentration of patent output has been mapped, clusters are identified using the heat map approach. 'Heat maps', use the Kernel Density Estimation (KDE) method (Parzen 1962; Rosenblatt 1956), a spatial interpolation technique that is frequently used to spatially compile data about crime levels, traffic accidents, property values as well as temperature, from which the 'heat map'-terminology originates. The KDE method appears to offer some important advantages over the cluster identification method used by Alcácer and Zhao (2016). Firstly, Alcácer and Zhao (2016) appears to have gone through a labour-intensive process of assigning patent addresses to particular cities and then combining cities that are in close proximity (less than 40 mi or 64.4 km) into the same cluster. The KDE method skips the need to assign an address to a city as the weightings of nearby locations are combined, thus neighbourhoods, neighbouring cities, adjascent villages or a university campus are automatically interpolated into one big 'hot spot'.

When applying the KDE method decisions must be made about two important variables: the interpolation range and the concentration threshold for recognizing an area as being of high concentration. The interpolation range can be decided based on several criteria, for example Van Egeraat et al. (2018) uses commuting distance while Alcácer and Zhao (2016) uses 20 mi (32.2 km), with no justification given. Acs et al. (2002) notes that within a 50 mi (80.5 km) distance from the boundaries of a metropolitan statistical area, there is still some positive innovation effect. The distance cited by Acs et al. (2002) is about four times the largest average daily commuting distance of a US city (Atlanta, GA, average commuting distance of 20.6 km) (Kneebone and Holmes 2015).

To identify a suitable interpolation range multiple distances are explored (10, 25 and 50 km). To illustrate the differences between interpolation ranges consider the region

² For details about TwoFishes, see: https://github.com/foursquare/fsqio (accessed 24 March 2019).



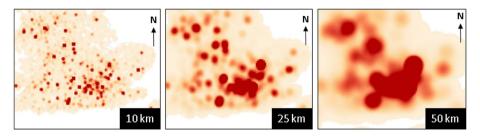


Fig. 1 Inventive activity heatmaps of London and environs with 10, 25 and 50 km radius

around London, United Kingdom, which is a high R&D city-region with multiple poles, including London, Cambridge and Oxford, which are also home to well-known universities. Figure 1 shows the heatmaps for the 10, 25 and 50 km interpolation ranges. A 10 or 25 km range identifies many small but spatially proximate high R&D regions. A 50 km range (right) identifies three large urban regions, from top to bottom: Manchester, Birmingham and London (with London incorporating Cambridge, Oxford and Southampton). The distance from In this research a distance of 50 km is chosen, which identifies high R&D regions at a spatial level that appears to coincide with broader urban agglomerations. The distance between central London and Cambridge, Oxford and Southampton is between 80 and 100 km, which is similar to the knowledge spillover distance Acs et al. (2002) determined to be significant.

To classify a region as part of an innovation city, the intensity of inventive activity must be within the 97.5th percentile. Using a lower threshold risks creating very large innovation regions which in some cases cover several countries, which would remove the urban and sub-national spatial scale of the innovation city. However even at the 97.5th percentile some very large innovation city-regions are identified, including those surrounding Tokyo, Frankfurt and New York. However because of the empirical support that 50 km is a suitable distance (interpolation range) at which intensive collaborations do occur, the threshold is not raised further, which would remove smaller cities from the sample.

The smallest innovation city-region that the chosen radius and threshold identify has an area of 112 km² (Sapporo, Japan) and the largest has an area of 100,260 km² (centered on Tokyo, Japan and encompassing Nagoya and Osaka). Although this large range in sizes challenge the traditional conception of what constitutes a city, there do not appear to be any strong theoretical reasons to exclude large city regions. Furthermore, when the city-region sizes are placed on a logarithmic scale, they are normally distributed.

Two more technical notes concerning the city identification process need to be made. First, in order to track the cities' development over time, city-regions are identified based on the total patenting activity during the entire study period, from 1997 to 2011, and therefore a constant city-region boundary is used. The constant boundary facilitates comparisons over time but they could remove some of the possible spatial dynamics from the research.

Second, city-regions are named based on the largest city located within the cluster boundaries. City names, coordinates and populations are obtained from the Simple Maps World Cities dataset which is based on public information from the National Geospatial Intelligence Agency (NGIA), United States Census Bureau, United States Geological



Table 1 R&D expenditure (2005 constant PPP million US\$) per patent grant, three periods, selected countries (sources: UNESCO Institute of Statistics and USPTO, 2018)

	1997–2001	2002-2006	2007–2011
China	78.5	38.9	29.9
Germany	4.5	5.2	6.5
Great Britain	5.7	7.1	8.2
Japan	2.7	2.9	3.3
Korea (Republic of)	4.0	2.7	3.9
Netherlands	6.0	6.0	6.9
United States	2.6	3.0	3.9

Survey and the National Aeronautics and Space Administration (NASA). Its free basic dataset contains information about approximately 13,000 large cities worldwide.³

Cities identified in an area with no significant population center are subject to additional scrutiny and often lead to the identification of miscoded locations (false positives), a problem that seems to occur in South Korea and Japan where 11 miscoded locations are identified, including Daejeon (South Korea), Yokkaichi, Kurashiki, Nara, Sendai, Kanagawa and Tochigi (Japan).

Patent activity heatmaps for Europe, the United States, Northeast Asia and Australia and New Zealand are included in the "Appendix".

3.2 R&D expenditure estimation

R&D expenditure is chosen as the indicator to compare global city-regions because it comes closest to describing the real volume of innovation activity that is taking place. Other options, such as directly comparing patent output, suffer from a number of downsides such as different patenting frequencies between sectors and countries, which can lead to over- and under-estimations of R&D activity. R&D expenditure is also a strong predictor of future economic growth (Marković et al. 2017).

R&D expenditure and patent output are closely correlated (Hagedoorn and Cloodt 2003; Lanjouw and Schankerman 2004; Milovančević et al. 2017; Squicciarini et al. 2013) and therefore patents are a suitable proxy for estimating R&D expenditure. International R&D expenditure statistics are obtained from the UNESCO Institute of Statistics who publish gross domestic expenditure on R&D (GERD) data for a number of countries starting in 1996.⁴

Unfortunately a number of biases and variances exist that influence patent data. Therefore a patent is not equal to a fixed amount of R&D expenditure in every country. In fact, the range is quite large as can be seen in Table 1.

One reason for this variation is the so-called 'home bias'. The home bias consists of two parts: the home country's patents are overrepresented in its national patent database and home country patents are cited more frequently than foreign patents (Bacchiocchi

⁴ Science, technology and innovation: Gross domestic expenditure on R&D (GERD), GERD as a percentage of GDP, GERD per capita and GERD per researcher statistics are available from: http://data.uis.unesco.org/ (accessed 24 March 2019).



³ The Simple Maps World Cities database can be downloaded at: https://simplemaps.com/data/world-cities (accessed 24 March 2019).

Table 2 Worldwide high R&D city-regions, top-20 (R&D expenditure in 2005 constant PPP US\$, world share)

Rank	1997–2001	2002–2006	2007–2011
1	Tokyo, JP (\$389b, 17.3%)	Tokyo, JP (\$468b, 16%)	Tokyo, JP (\$572b, 14%)
2	San Francisco, US (\$150b, 6.7%)	San Francisco, US (\$202b, 6.9%)	San Francisco, US (\$288b, 7%)
3	Frankfurt, DE (\$134b, 5.9%)	Frankfurt, DE (\$159b, 5.4%)	Seoul, KR (\$203b, 5%)
4	New York, US (\$131b, 5.8%)	New York, US (\$135b, 4.6%)	Frankfurt, DE (\$203b, 5%)
5	Los Angeles, US (\$91b, 4.1%)	Los Angeles, US (\$119b, 4.1%)	New York, US (\$172b, 4.2%)
6	Seoul, KR (\$63b, 2.8%)	Seoul, KR (\$98b, 3.3%)	Los Angeles, US (\$167b, 4.1%)
7	Boston, US (\$60b, 2.7%)	Taipei, TW (\$81b, 2.7%)	Beijing, CN (\$166b, 4.1%)
8	Taipei, TW (\$54b, 2.4%)	Boston, US (\$75b, 2.5%)	Shanghai, CN (\$105b, 2.6%)
9	Paris, FR (\$53b, 2.3%)	Beijing, CN (\$65b, 2.2%)	Taipei, TW (\$101b, 2.5%)
10	Chicago, US (\$52b, 2.3%)	London, GB (\$61b, 2.1%)	Boston, US (\$98b, 2.4%)
11	London, GB (\$48b, 2.1%)	Chicago, US (\$59b, 2%)	Hong Kong, HK (\$88b, 2.1%)
12	The Hague, NL (\$37b, 1.6%)	Seattle, US (\$58b, 2%)	Seattle, US (\$84b, 2.1%)
13	Detroit, US (\$36b, 1.6%)	Paris, FR (\$55b, 1.9%)	London, GB (\$77b, 1.9%)
14	Minneapolis, US (\$34b, 1.5%)	Moscow, RU (\$47b, 1.6%)	Chicago, US (\$75b, 1.8%)
15	Moscow, RU (\$32b, 1.4%)	The Hague, NL (\$46b, 1.6%)	Sydney, AU (\$69b, 1.7%)
16	Beijing, CN (\$31b, 1.4%)	Shanghai, CN (\$46b, 1.6%)	Moscow, RU (\$65b, 1.6%)
17	Essen, DE (\$30b, 1.3%)	Detroit, US (\$43b, 1.5%)	Paris, FR (\$61b, 1.5%)
18	Washington, D.C., US (\$29b, 1.3%)	Munich, DE (\$42b, 1.4%)	Detroit, US (\$57b, 1.4%)
19	Munich, DE (\$29b, 1.3%)	Sydney, AU (\$41b, 1.4%)	The Hague, NL (\$53b, 1.3%)
20	Seattle, US (\$27b, 1.2%)	Minneapolis, US (\$41b, 1.4%)	Munich, DE (\$51b, 1.3%)

and Montobbio 2010; de la Potterie and De Rassenfosse 2008). Other causes of variation in national patent output relative to R&D expenditure include differences in a country's sectoral R&D profile, as patenting frequencies vary between sectors (Breschi et al. 2000; Kleinknecht et al. 2002; Toivanen and Suominen 2015) and differences in national institutions and laws (De Rassenfosse and de la Potterie 2009).

Therefore the R&D expenditure (EXP) in a particular city-region (C) is estimated using the number of patents generated within the city-region (PAT_C) multiplied by the R&D expenditure per patent (EPP) of the nation or territory (N) in which the city-region is primarily located, as shown in the equation below:

$$EXP_C = EPP_N \times PAT_C$$

4 Results

The identification process of high R&D city-regions during three periods (1997–2001, 2002–2006, 2007–2011) provides a global and historical overview of the real worldwide spatial distribution of R&D expenditure. A total of 132 city-regions are identified of which 76 are in North America, 36 are in Europe and 19 are in Asia-Pacific (20 if Israel is included). No high R&D city-regions are identified in South America, Africa or India. An overview of the top-20 city-regions is provided in Table 2.



Rank	1997–2001	2002–2006	2007–2011
1	Tokyo, JP (\$389b, 17.3%)	Tokyo, JP (\$468b, 16%)	Tokyo, JP (\$572b, 14%)
2	Seoul, KR (\$63b, 2.8%)	Seoul, KR (\$98b, 3.3%)	Seoul, KR (\$203b, 5%)
3	Taipei, TW (\$54b, 2.4%)	Taipei, TW (\$81b, 2.7%)	Beijing, CN (\$166b, 4.1%)
4	Beijing, CN (\$31b, 1.4%)	Beijing, CN (\$65b, 2.2%)	Shanghai, CN (\$105b, 2.6%)
5	Singapore, SG (\$14b, 0.6%)	Shanghai, CN (\$46b, 1.6%)	Taipei, TW (\$101b, 2.5%)
6	Shanghai, CN (\$13b, 0.6%)	Sydney, AU (\$41b, 1.4%)	Hong Kong, HK (\$88b, 2.1%)
7	Sydney, AU (\$11b, 0.5%)	Singapore, SG (\$23b, 0.8%)	Sydney, AU (\$69b, 1.7%)
8	Kaohsiung, TW (\$9b, 0.4%)	Hong Kong, HK (\$19b, 0.7%)	Singapore, SG (\$37b, 0.9%)
9	Hong Kong, HK (\$5b, 0.2%)	Kaohsiung, TW (\$14b, 0.5%)	Kaohsiung, TW (\$33b, 0.8%)
10	Melbourne, AU (\$5b, 0.2%)	Melbourne, AU (\$9b, 0.3%)	Melbourne, AU (\$18b, 0.5%)

Table 3 Asia-Pacific high R&D city-regions, top-10 (R&D expenditure in 2005 constant PPP US\$, world share)

At a global level the rising R&D expenditure of city-regions in Asia is clearly noticeable among the top-20 city-regions. Seoul, followed by Beijing, Shanghai, Taipei and Hong Kong all experience an increase in R&D expenditure and ranking position. During the 1997–2001 period 3 Asian city-regions are in the top 10. By 2007–2011 Shanghai and Beijing have also entered the top 10, making half of the top 10 R&D city-regions from Asia. City-regions in other parts of the world also experience higher R&D expenditure, but many have a constant or falling share of global R&D expenditure and rank due to the faster growth of R&D expenditure in Asian city-regions.

Top-20 city-regions which experience a fall in world share during the 1997–2011 period include Tokyo, Frankfurt, New York, Boston, Paris, Chicago, London, The Hague, Detroit, Minneapolis, Essen and Washington DC. City-regions outside Asia which increased their share are San Francisco, Los Angeles, Seattle, Moscow and Sydney.

The geography of R&D expenditure growth and relative decline appears to be defined by rising R&D expenditure in Asia and in city-regions along the Pacific Coast of the United States, Australia (Sydney) and Russia (Moscow). Declining top-20 city-regions, in relative terms, are found in Western Europe, Japan and the Northeastern United States.

Despite these changes the largest R&D city-regions are Tokyo and San Francisco during all three periods, with Tokyo experienceing a relative decline and San Francisco growing both in absolute and relative terms.

While the rise of Asian city-regions provides evidence of a global shift, the rise of Asian cluster-regions is partly due to the high concentration of R&D expenditure within Asian countries (Crescenzi and Rodríguez-Pose 2017; Yoon and Park 2017). Of the 19 Asia Pacific city regions in the study 5 are in the top-10 by 2006–2011 (26%) compared to 4 of 76 North American city-regions (5%) and 1 of 36 European city-regions (3%).

To some extent this result can be attributed to the way in which city-regions are identified, which is based on inventor-weighted patent data. Due to the home bias effect and other variances mentioned in the Methodology this may lead to an over-detection of city-regions in the United States (favourable home bias effect) and an under-detection of city-regions in a country such as China, where R&D expenditure per USPTO patent grant is very high (see Table 1, earlier). However when the city-region threshold is raised to a minimum R&D expenditure of \$10b in 2007–2011, 5 of 10 Asian city-regions appear in the top-10 (50%) compared to 4 of 29 North American city-regions (14%) and 1 of 19



Rank	1997–2001	2002–2006	2007–2011
1	Frankfurt, DE (\$134b, 5.9%)	Frankfurt, DE (\$159b, 5.4%)	Frankfurt, DE (\$203b, 5%)
2	Paris, FR (\$53b, 2.3%)	London, GB (\$61b, 2.1%)	London, GB (\$77b, 1.9%)
3	London, GB (\$48b, 2.1%)	Paris, FR (\$55b, 1.9%)	Moscow, RU (\$65b, 1.6%)
4	The Hague, NL (\$37b, 1.6%)	Moscow, RU (\$47b, 1.6%)	Paris, FR (\$61b, 1.5%)
5	Moscow, RU (\$32b, 1.4%)	The Hague, NL (\$46b, 1.6%)	The Hague, NL (\$53b, 1.3%)
6	Essen, DE (\$30b, 1.3%)	Munich, DE (\$42b, 1.4%)	Munich, DE (\$51b, 1.3%)
7	Munich, DE (\$29b, 1.3%)	Essen, DE (\$30b, 1%)	Essen, DE (\$36b, 0.9%)
8	Milan, IT (\$25b, 1.1%)	Milan, IT (\$28b, 0.9%)	Barcelona, ES (\$35b, 0.9%)
9	København, DK (\$16b, 0.7%)	Barcelona, ES (\$26b, 0.9%)	Milan, IT (\$32b, 0.8%)
10	Stockholm, SE (\$14b, 0.6%)	København, DK (\$19b, 0.6%)	København, DK (\$24b, 0.6%)

Table 4 European high R&D city-regions, top-10 (R&D expenditure in 2005 constant PPP US\$, world share)

European city-regions (5%). So the observation of a high spatial concentration of R&D expenditure in a small number of city-regions within Asia Pacific holds true.

4.1 Asia-Pacific

Being home to city-regions with fast-growing R&D expenditure, developments taking place in the Asia-Pacific region are of particular interest. Table 3 shows the R&D expenditure of the top-10 city-regions in Asia-Pacific.

Among the top-10 city-regions, all regions experience increasing R&D expenditure and an increase in their global share with the notable exception of Tokyo and Taipei, whose global share is declining during the 1997–2011 period. Among the top-10 clusters two notable groups of city-regions can be identified: city-regions that are rising in ranking position and those that are falling in ranking position, often despite increasing their world share. Rising city-regions are Beijing, Shanghai and Hong Kong. Falling city-regions, in relative terms, are Taipei, Singapore and Kaohsiung. Tokyo should also be considered a falling city-region as it is losing global share but it maintains its first position within Asia-Pacific (and globally).

Based on these developments there appears to be a shift of R&D expenditure towards the 3 Chinese city-regions, with Beijing and Shanghai being closely followed by Hong Kong. If this growth trajectory continues it is concievable that Beijing will overtake Seoul during the next decade as the difference in R&D expenditure between them is about 22% (\$37b) in 2007–2011, having fallen from 51% (\$33b) during the previous 2002–2006. The gap between Seoul and Tokyo continues to be significant at 181% (\$369b), so Tokyo may well remain the largest high R&D city-region for the foreseeable future.

While the rise of Chinese city-regions is significant, there is a notable lack of second-tier Chinese city-regions. While there may be methodological reasons for the failure to detect smaller Chinese high R&D city-regions, Korea, Japan and Taiwan all have a number of second-tier city-regions such as Kaohsiung (Taiwan), Busan, Daegu (Korea), Fukuoka and Hiroshima (Japan). While the large concentrations in Tokyo and Seoul suggest that the three largest Chinese city-regions could grow much further, a second tier of high R&D city-regions may also develop.



Rank	1997–2001	2002–2006	2007–2011
1	San Francisco, US (\$150b, 6.7%)	San Francisco, US (\$202b, 6.9%)	San Francisco, US (\$288b, 7%)
2	New York, US (\$131b, 5.8%)	New York, US (\$135b, 4.6%)	New York, US (\$172b, 4.2%)
3	Los Angeles, US (\$91b, 4.1%)	Los Angeles, US (\$119b, 4.1%)	Los Angeles, US (\$167b, 4.1%)
4	Boston, US (\$60b, 2.7%)	Boston, US (\$75b, 2.5%)	Boston, US (\$98b, 2.4%)
5	Chicago, US (\$52b, 2.3%)	Chicago, US (\$59b, 2%)	Seattle, US (\$84b, 2.1%)
6	Detroit, US (\$36b, 1.6%)	Seattle, US (\$58b, 2%)	Chicago, US (\$75b, 1.8%)
7	Minneapolis, US (\$34b, 1.5%)	Detroit, US (\$43b, 1.5%)	Detroit, US (\$57b, 1.4%)
8	Washington DC, US (\$29b, 1.3%)	Minneapolis, US (\$41b, 1.4%)	Minneapolis, US (\$51b, 1.3%)

Toronto, CA (\$36b, 1.2%)

Washington DC, US (\$32b,

1.1%)

Toronto, CA (\$47b, 1.2%)

Houston, US (\$45b, 1.1%)

Table 5 North American high R&D city-regions, top-10 (R&D expenditure in 2005 constant PPP US\$, world share)

4.2 Europe

Seattle, US (\$27b, 1.2%)

Dallas, US (\$27b, 1.2%)

9

10

Viewed globally, European high R&D city regions are in relative decline, being overtaken by large city-regions in Asia, even though in absolute terms R&D expenditure in European city-regions is increasing. Table 4 shows the R&D expenditure of the top-10 city-regions in Asia-Pacific.

Within Europe a number of city-regions are increasing their relative position, notably London, Moscow, Munich and Barcelona. Declining in ranking position are Paris, The Hague, Essen and Milan. Frankfurt, which maintains its top position, is declining in terms of its world share. However of the rising city-regions only Moscow and Barcelona are increasing their share of world R&D expenditure to 0.9% and 1.6% respectively. One of the most significant downward movements is by Paris, from 2nd position to 4th and reducing its global share from 2.3 to 1.5%.

Most top-10 city-regions in Europe show only very low growth in R&D expenditure, especially when compared to other parts of the world. Within the European Union Barcelona is the only top-10 city-region with a significant rise in R&D expenditure that exceeds or is in line with the global level of growth. Therefore in relative terms, Europe's global position in R&D is likely to continue its decline.

4.3 North America

North America is the region in which the largest number (76) of high R&D city-regions is identified, and the 10 largest regions for each period are shown in Table 5.

As in other regions of the world, R&D expenditure in the top-10 city-regions is growing although only two city-regions, Seattle and San Francisco, are experiencing an increase in their global share. Los Angeles and Toronto maintain their global share by growing at close to the global average rate.

As noted earlier city-regions located on the Pacific coast appear to be growing at a faster rate than those positioned along the Atlantic coast (Boston, New York, Washington) and



in the Midwest (Chicago, Detroit and Minneapolis). Hence the United States appears to be experiencing a domestic shift in terms of R&D towards the Pacific coast. If current trends persist it is likely that Los Angeles will overtake New York as the second-largest R&D city-region in North America during the coming decade.

5 Discussion

The identifying of high R&D city-regions based on patent data provides an interesting subnational perspective on changes in the spatial distribution of global innovation activity, and at the same time raises some important technical and theoretical questions.

From a technical perspective, the significant differences in R&D expenditure per USPTO patent grant (Table 1) is a cause for concern as it calls into question the validity of identifying high R&D city-regions based on patent data. In countries with relatively high R&D expenditure per patent such as China, smaller city-regions may not be identified because of the low number of patents. On the other hand too many clusters may be identified in the United States and Japan where R&D expenditure per patent is relatively low.

A second concern is the potential underidentification of city-regions in emerging economies because city-regions are identified using 15 years of accumulated patent data from 1997 to 2011. During the earlier part of that period patenting in emerging economies may have been relatively low, thus lowering the cumulative patenting output based on which city-regions are identified. However during the later part of the period patenting activity was high as these city-regions experience rapid R&D expenditure growth. Therefore identifying city-regions based only on relatively recent data may be another way to avoid underidentification in emerging economies.

Based on these technical concerns the methodology is likely to have failed to identify smaller city-regions in countries with high R&D expenditure per patent such as China (\$78 m), Russia (\$45 m), Spain (\$21 m). Nevertheless city-regions were successfully identified in these countries using the current methodology in these countries and these city-regions feature among the largest in the world.

From a theoretical perspective, the identification of large R&D intensive city-regions is an interesting phenomenon. Are Tokyo-Nagoya-Osaka or Frankfurt-Zurich-Grenoble or New York-Philadelphia really part of a single interconnected city-region or corridor? Or are they a set of spatially proximate but essentially separate cities? Regardless of the answer, the existence of large spatial concentrations of R&D is a notable geographical phenomenon that may be due more to urbanization in general than due to R&D activity specifically.

The existence of (very) large city-regions alongside many smaller city-regions also raises questions about the inequality of R&D activity between major cities around the world. While inequality in research output can be measured in many different ways (Jeon and Kim 2018), the decline of some large city-regions and the growth of others in terms of their R&D expenditure has implications for the global distribution of economic activity and prosperity. The ranking tables in this paper clearly show that in terms of R&D expenditure, the world is highly unequal. The tables also show that city-regions in Asia-Pacific such as Seoul, Beijing, Shanghai and Hong Kong, are showing positive path-creation as they rapidly improve their global position.

Identifying innovation city-regions using patent data also opens the door to further analysis of their development, including changes in the patent co-invention networks or Triple



Helix collaboration networks (Leydesdorff and Persson 2010; Leydesdorff and Sun 2009), patent inventor-assignee networks (Bhattacharya 2004), patent quality indicators (Gautam et al. 2014; Hagedoorn and Cloodt 2003; Lanjouw and Schankerman 2004) or changes in the kinds of collaborations and research taking place within a city-region's innovation system (Park and Leydesdorff 2010; Stek and Van Geenhuizen 2015), including the distance of a city-region from the technological frontier (Toivanen and Suominen 2015) and which important technologies are being researched (Jung 2017). These are all avenues that offer opportunities to gain further insight into the sub-national spatial dynamics of global R&D expenditure and innovation activity.

6 Conclusion

This paper offers a methodology for mapping high R&D city-regions worldwide and its results provide meaningful insight into the changing global distribution of innovation activity, which is shifting towards a number of cities in Asia and three Chinese cities in particular.

The relatively stable position of the largest high R&D city-regions suggest there is considerable path-dependence in the global research and innovation system. However the rise of certain cities in Asia suggests path-creating factors are present there, either due to public policy interventions, foreign investment in R&D, or more likely, a combination of the two (Crescenzi and Jaax 2017; Dicken 2007).

In addition to Asian cities, within Europe and North America there are a number of city-regions with R&D growth that exceeds the growth rate of the broader region in which they are located. These high-growth city-regions include San Francisco, Seattle, Barcelona, London, Moscow and Munich and their innovation development may be of interest to researchers seeking to understand the path-creating factors that influence local R&D activity. As highlighted in the discussion, patent data can also be used to derive other indicators related to networks, organizations and technology, which may provide further insight into these city-regions.

Aside from providing insight into the global geography of R&D, the research also demonstrates some of the challenges of estimating R&D expenditure using patent data. The identification of very large city-regions also raises a notable theoretical question about the existence of very large city-regions which include multiple cities. Do existing agglomeration theories apply to such multi-city 'meta regions'? Or should the methodology be adapted? These are all questions that can hopefully be addressed in future research as well.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Figs. 2, 3, 4 and 5.



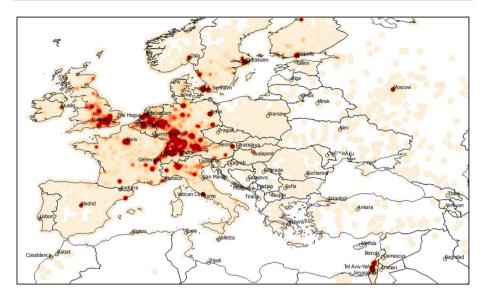


Fig. 2 Inventive activity heatmap of European region

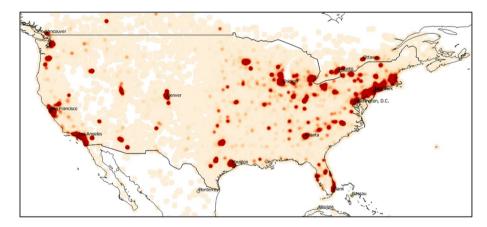


Fig. 3 Inventive activity heatmap of the continental United States



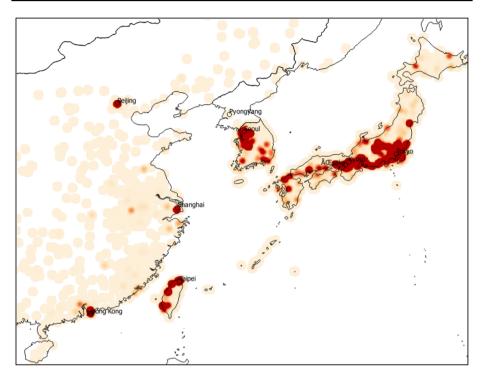


Fig. 4 Inventive activity heatmap of Northeast Asia

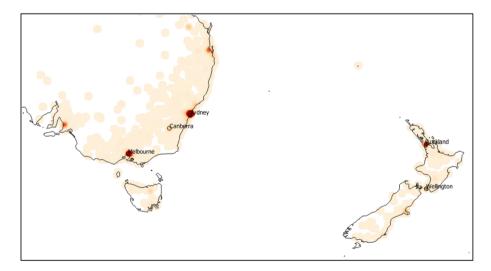


Fig. 5 Inventive activity heatmap of southeast Australia and main islands of New Zealand

References

- Acs, Z.J., Anselin, L., Varga, A.: Patents and innovation counts as measures of regional production of new knowledge. Res. Policy 31(7), 1069–1085 (2002)
- Alcácer, J., Zhao, M.: Zooming in: a practical manual for identifying geographic clusters. Strateg. Manag. J. 37(1), 10–21 (2016)
- Audretsch, D.B., Lehmann, E.E., Wright, M.: Technology transfer in a global economy. J. Technol. Transf. **39**(3), 301–312 (2014)
- Bacchiocchi, E., Montobbio, F.: International knowledge diffusion and home-bias effect: do uspto and epo patent citations tell the same story? Scand. J. Econ. 112(3), 441–470 (2010)
- Bathelt, H., Malmberg, A., Maskell, P.: Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. Prog. Hum. Geogr. 28(1), 31–56 (2004)
- Beaverstock, J.V., Smith, R.G., Taylor, P.J.: A roster of world cities. Cities 16(6), 445–458 (1999)
- Bhattacharya, S.: Mapping inventive activity and technological change through patent analysis: a case study of India and China. Scientometrics **61**(3), 361–381 (2004)
- Binz, C., Truffer, B.: Global innovation systems: a conceptual framework for innovation dynamics in transnational contexts. Res. Policy 46(7), 1284–1298 (2017)
- Boschma, R.A., Frenken, K.: Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. J. Econ. Geogr. 6, 273–302 (2006)
- Breschi, S., Malerba, F., Orsenigo, L.: Technological regimes and schumpeterian patterns of innovation. Econ. J. 110(463), 388–410 (2000)
- Capello, R.: Spatial spillovers and regional growth: a cognitive approach. Eur. Plan. Stud. 17(5), 639–658 (2009)
- Castells, M.: Silicon valley: where it all began. Technopoles of the World, pp. 22–38. Routledge, London (2014)
- Crescenzi, R., Jaax, A.: Innovation in Russia: the territorial dimension. Econ. Geogr. 93(1), 66-88 (2017)
- Crescenzi, R., Rodríguez-Pose, A.: Innovation and Regional Growth in the European Union. Springer, Heidelberg (2011)
- Crescenzi, R., Rodríguez-Pose, A.: The geography of innovation in China and India. Int. J. Urban Reg. Res. 41(6), 1010–1027 (2017)
- de la Potterie, B.P., De Rassenfosse, G.: Policymakers and the R&D-patent relationship. Intereconomics 43(6), 377–380 (2008)
- De Rassenfosse, G., de la Potterie, B.V.: A policy insight into the R&D-patent relationship. Res. Policy 38(5), 779-792 (2009)
- Dicken, P.: Global Shift: Mapping the Changing Contours of the World Economy. SAGE Publications, London (2007)
- Duranton, G., Overman, H.G.: Testing for localization using micro-geographic data. Rev. Econ. Stud. 72(4), 1077–1106 (2005)
- Dutta, S., Lanvin, B.: The Global Innovation Index 2013: The Local Dynamics of Innovation. Cornell University, Ithaca (2016)
- Feldman, M.P.: The Geography of Innovation, vol. 2. Kluwer Academic Publishers, Dordrecht (1994)
- Feldman, M. P., Kogler, D. F.: Stylized facts in the geography of innovation. In: Handbook of the Economics of Innovation, vol. 1, pp. 381–410. Elsevier, Amsterdam (2010)
- Frenken, K., Cefis, E., Stam, E.: Industrial dynamics and clusters: a survey. Reg. Stud. 49(1), 10-27 (2015)
- Fujita, M., Krugman, P.R., Venables, A.J.: The Spatial Economy: Cities, Regions, and International Trade. MIT Press, Cambridge (2001)
- Gautam, P., Kodama, K., Enomoto, K.: Joint bibliometric analysis of patents and scholarly publications from cross-disciplinary projects: implications for development of evaluative metrics. J. Contemp. East. Asia 13(1), 19–37 (2014)
- Gertler, M.S.: Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). J. Econ. Geogr. 3(1), 75–99 (2003)
- Gertler, M.S., Wolfe, D.A.: Spaces of knowledge flows: clusters in a global context. In: Asheim, B., Cooke, P., Martin, R. (eds.) Clusters and Regional Development, pp. 218–235. Routledge, London (2006)
- Grillitsch, M., Nilsson, M.: Firm performance in the periphery: on the relation between firm-internal knowledge and local knowledge spillovers. Reg. Stud. **51**(8), 1219–1231 (2017)
- Hagedoorn, J., Cloodt, M.: Measuring innovative performance: is there an advantage in using multiple indicators? Res. Policy **32**(8), 1365–1379 (2003)
- Hamstead, Z.A., Fisher, D., Ilieva, R.T., Wood, S.A., McPhearson, T., Kremer, P.: Geolocated social media as a rapid indicator of park visitation and equitable park access. Comput. Environ. Urban Syst. 72, 38–50 (2018)



- Jeon, J., Kim, S.Y.: Is the gap widening among universities? On research output inequality and its measurement in the Korean higher education system. Qual. Quant. 52(2), 589–606 (2018)
- Jung, Y.: A semantic annotation framework for scientific publications. Qual. Quant. 51(3), 1009–1025 (2017)
- Kim, J., Lee, S.: Patent databases for innovation studies: a comparative analysis of USPTO, EPO, JPO and KIPO. Technol. Forecast. Soc. Chang. **92**, 332–345 (2015)
- Kleinknecht, A., Van Montfort, K., Brouwer, E.: The non-trivial choice between innovation indicators. Econ. Innov. New Technol. 11(2), 109–121 (2002)
- Kneebone, E., Holmes, N.: The Growing Distance Between People and Jobs in Metropolitan America. The Brookings Institution, Washington, DC (2015)
- Lanjouw, J.O., Schankerman, M.: Patent quality and research productivity: measuring innovation with multiple indicators. Econ. J. 114(495), 441–465 (2004)
- Lee, K., Lim, C.: Technological regimes, catching-up and leapfrogging: findings from the Korean industries. Res. Policy **30**(3), 459–483 (2001)
- Lee, Y.-S., Tee, Y.-C., Kim, D.-W.: Endogenous versus exogenous development: a comparative study of biotechnology industry cluster policies in South Korea and Singapore. Environ. Plan. C Gov. Policy 27(4), 612–631 (2009)
- Leydesdorff, L., Persson, O.: Mapping the geography of science: distribution patterns and networks of relations among cities and institutes. J. Am. Soc. Inform. Sci. Technol. **61**(8), 1622–1634 (2010)
- Leydesdorff, L., Sun, Y.: National and international dimensions of the triple helix in Japan: university—industry—government versus international coauthorship relations. J. Am. Soc. Inform. Sci. Technol. **60**(4), 778–788 (2009)
- Leydesdorff, L., Alkemade, F., Heimeriks, G., Hoekstra, R.: Geographic and technological perspectives on "photovoltaic cells:" patents as instruments for exploring innovation dynamics. Internetquelle: arXiv: 1401.2778 (2014)
- Locke, R.M., Wellhausen, R.L.: Production in the Innovation Economy. MIT Press, Cambridge (2014)
- Malecki, E.J.: The geography of innovation. In: Fischer, M., Nijkamp, P. (eds.) Handbook of Regional Science, pp. 375–389. Springer, Berlin (2014)
- Marković, D., Mladenović, I., Milovančević, M.: Estimation of the most influential science and technology factors for economic growth forecasting by soft computing technique. Qual. Quant. 51(3), 1133–1146 (2017)
- Martin, R., Sunley, P.: Deconstructing clusters: chaotic concept or policy panacea? J. Econ. Geogr. 3(1), 5–35 (2003)
- Maskell, P., Bathelt, H., Malmberg, A.: Building global knowledge pipelines: the role of temporary clusters. Eur. Plan. Stud. **14**(8), 997–1013 (2006)
- Milovančević, M., Marković, D., Nikolić, V., Mladenović, I.: Determination of the most influential factors for number of patents prediction by adaptive neuro-fuzzy technique. Qual. Quant. **51**(3), 1207–1216 (2017)
- Park, H.W., Leydesdorff, L.: Longitudinal trends in networks of university-industry-government relations in South Korea: the role of programmatic incentives. Res. Policy **39**(5), 640–649 (2010)
- Parzen, E.: On estimation of a probability density function and mode. Ann. Math. Stat. 33(3), 1065–1076 (1962)
- Porter, M.E.: Clusters and the new economics of competition. Harv. Bus. Rev. 76(6), 77–90 (1998)
- Rosenblatt, M.: Remarks on some nonparametric estimates of a density function. Ann. Math. Stat. 27, 832–837 (1956)
- Saxenian, A.: Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University Press, Boston (1996)
- Schwab, K., Sala-i-Martin, X.: The Global Competitiveness Report 2015–2016. World Economic Forum, Geneva (2015)
- Sessions, C., Wood, S.A., Rabotyagov, S., Fisher, D.M.: Measuring recreational visitation at us national parks with crowd-sourced photographs. J. Environ. Manage. 183, 703–711 (2016)
- Sloan, L., Quan-Haase, A.: The Sage Handbook of Social Media Research Methods. Sage, London (2017)
- Squicciarini, M., Dernis, H., Criscuolo, C.: Measuring patent quality: Indicators of technological and economic value. Organisation for Economic Cooperation; Development (OECD), Paris (2013)
- Stek, P.E., Van Geenhuizen, M.S.: Measuring the dynamics of an innovation system using patent data: a case study of South Korea, 2001–2010. Qual. Quant. **49**(4), 1325–1343 (2015)
- Storper, M.: The Regional World: Territorial Development in a Global Economy. Guilford Press, New York (1997)
- Toivanen, H., Suominen, A.: The global inventor gap: distribution and equality of world-wide inventive effort, 1990–2010. PLoS ONE 10(4), e0122098 (2015)



- Van Egeraat, C., Morgenroth, E., Kroes, R., Curran, D., Gleeson, J.: A measure for identifying substantial geographic concentrations. Pap. Region. Sci. 97(2), 281–300 (2018)
- Yoon, J., Park, H.W.: Triple helix dynamics of South Korea's innovation system: a network analysis of interregional technological collaborations. Qual. Quant. 51(3), 989–1007 (2017)

