



PATENT DATA FOR INDUSTRY AND MARKET APPLICATIONS

Mariagrazia Squicciarini

OECD Directorate for Science, Technology and Innovation (STI)
Economic Analysis and Statistics Division (EAS)

mariagrazia.squicciarini@oecd.org

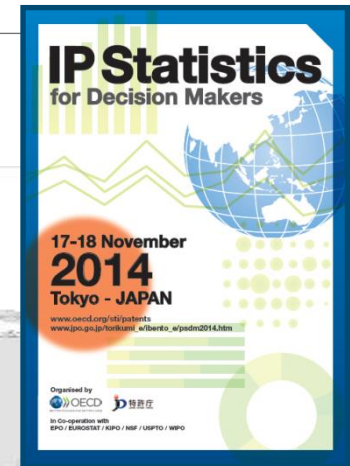
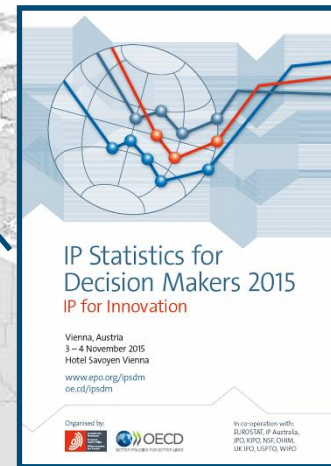


Intellectual Property data at the OECD

- In **2000**, OECD work initiated with patent data.
- **OECD Task Force on Patent Statistics**, collaboration with patent offices worldwide & inter governmental institutions
→ develop **statistical infrastructures** (e.g. PATSTAT) and **methodologies** (guidelines).
- Since **2007**, integration of other IP (Trademarks, Designs)
- In 2014, the Task Force became the **IP Statistics Task Force**
- Annual [conferences](#) on **IP Statistics for Decision Makers**



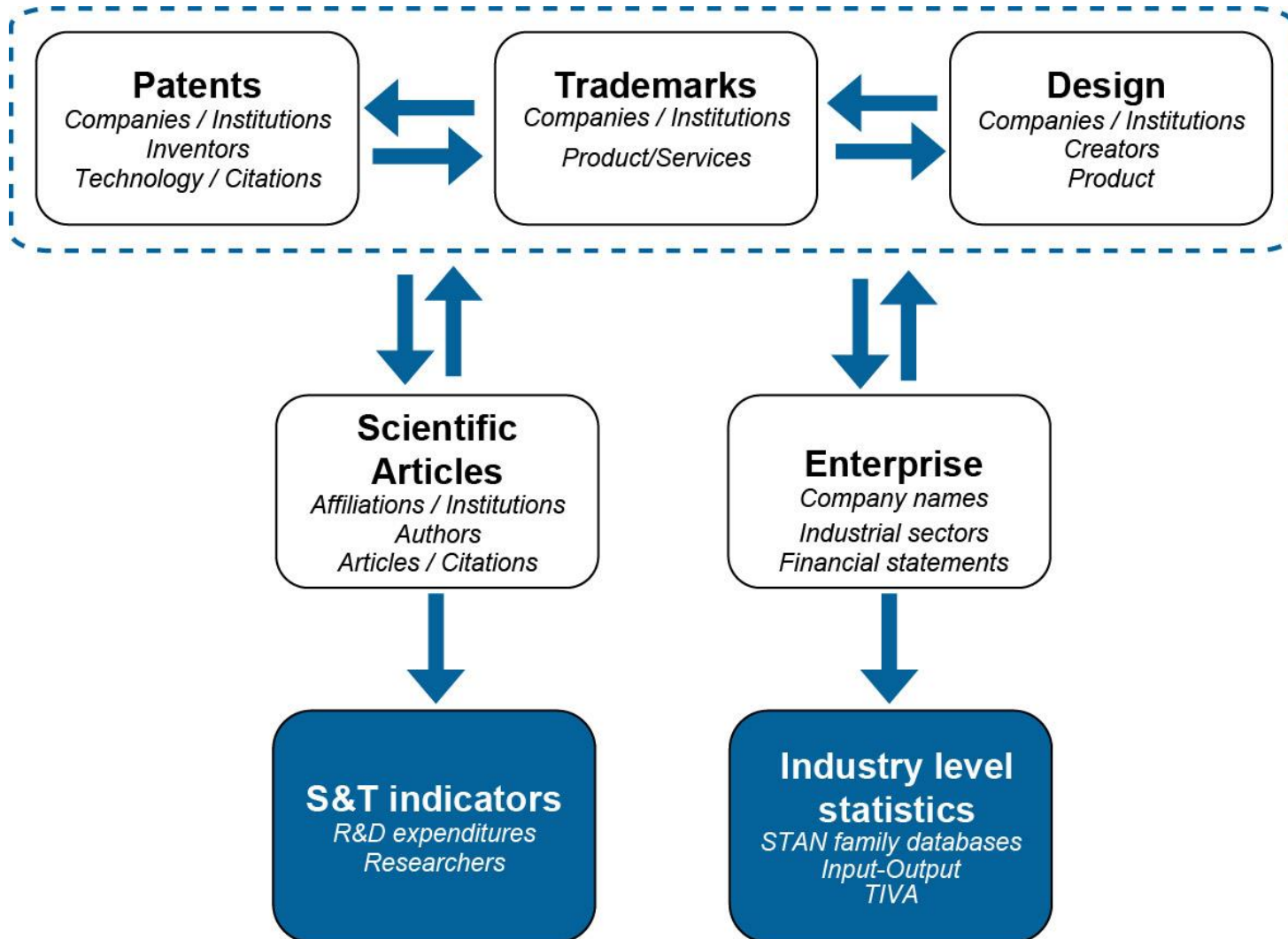
IP Statistics for Decision Makers 2016





STI Microdata lab

data infrastructure





AVAILABLE DATA



Patent Data

Raw data



Data collection

- **PATSTAT** (EPO's Worldwide Patent Statistical Database)
- **EPO Bibliographic data** (*for latest data on EPO and PCT filings*)
- **INPADOC** Legal Status Data & European Patent Register

Derived OECD patent datasets available upon request

- **Patent families:** Triadic, IP5, Equivalents
- **REGPAT:** patents at regional level (*EPO & PCT only*)
- **Patent citations:** consolidated patent citation data (*EPO, PCT, USPTO*)
- **Patent quality:** Indicators of patent value (*EPO, PCT, USPTO*)
- **HAN:** Harmonised applicant names (*based on matching to ORBIS*)

Series of pre-defined indicators available in **OECD.Stat**



Patent Families

Concept and definitions

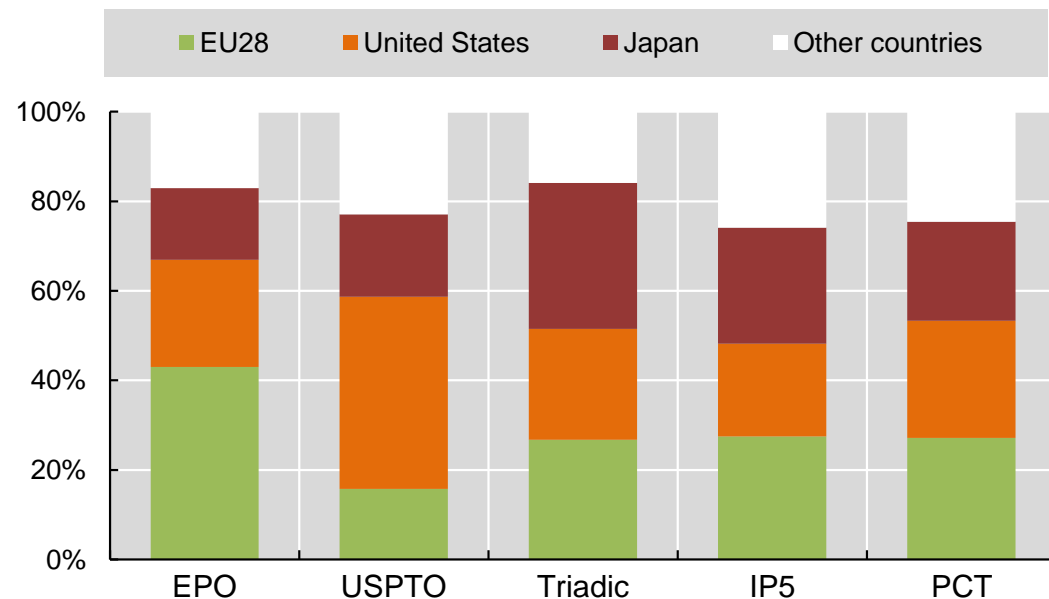
Patent families

- Patent protection is territorial: a patent application needs to be filed in each patent office where protection is sought.
- The first patent filing made to protect the invention may be followed by a series of subsequent filings, and all are considered as members of the same patent family.

- **Triadic Patent families :**
EPO + JPO + USPTO
area where most of R&D performed within OECD
Raw data available upon request

- **IP5 Patent families :**
EPO / JPO / KIPO / USPTO / SIPO
Families composed of 2 filings worldwide, and at least 1 patent filed in one of the Five IP offices

***Share of countries in patents,
national offices and family definitions, 2011***





Patents by Regions

The OECD REGPAT database

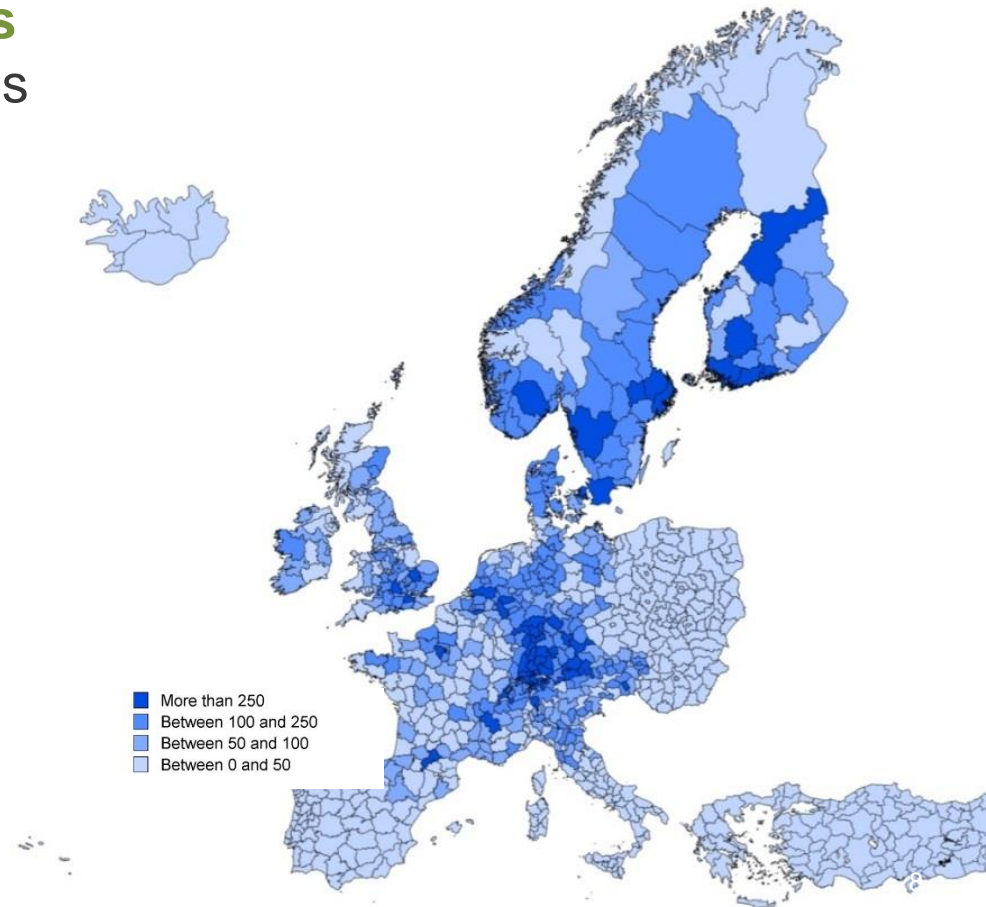
Lower unit of analysis - NUTS3/TL3 levels

- **Addresses** of inventors & applicants
- **String comparison algorithms** using country level lookup tables (postal codes / town names) to allocate regional code

- **5 500** regions from **40+** countries
OECD/EU28/BRICS
- Only for **EPO** & **PCT** patents

Raw data (OECD REGPAT)
available upon request

Patent intensity – PCT per capita





Measuring Patent Quality

Indicators

For each patent, try to capture:

- The **technological importance** of the invention
- its **economic value**
- and the possible **impact on subsequent technological developments**

12 indicators derived from literature, only based on **information contained** in patent documents:

- can be constructed for all patents,
- rely on a homogeneous set of information comparable across countries and over time.
- *Caveat: no information about e.g. market transactions or real use of the (patented) technology available.*

Detailed methodology : [OECD STI Working Papers 2013/03](#)
(including SQL code for building indicators)

Raw data available upon request (for EPO and USPTO)



Measuring Patent Quality

Indicators

- 1. Patent scope**
Number of IPC subclasses the invention is allocated to.
- 2. Patent family size**
Number of patent offices at which an invention has been protected by a patent.
- 3. Number of claims**
Number of claims per patent.
- 4. Background citations**
Number of patents cited in the patent document.
- 5. References to Non-Patent Literature**
Share of NPL citations in a patent document.
- 6. Forward citations**
Number of citations received in 5 year time after publication. Corrected for patent equivalents.
- 7. Breakthrough inventions**
Top 1% cited patents up to 5 years after publication.
- 8. Grant Lag**
Inverse function of the time elapsed between application and grant dates.
- 9. Generality index**
number and distribution of forward citations and IPC classes of the patents these citations come from.
- 10. Originality index**
same as generality – but based on backward citations.
- 11. Radicalness index**
number of IPC classes in which the patents cited by a given patent are, but in which the patent itself is not classified.
- 12. Patent renewal**
number of years during which a granted patent has been kept alive.

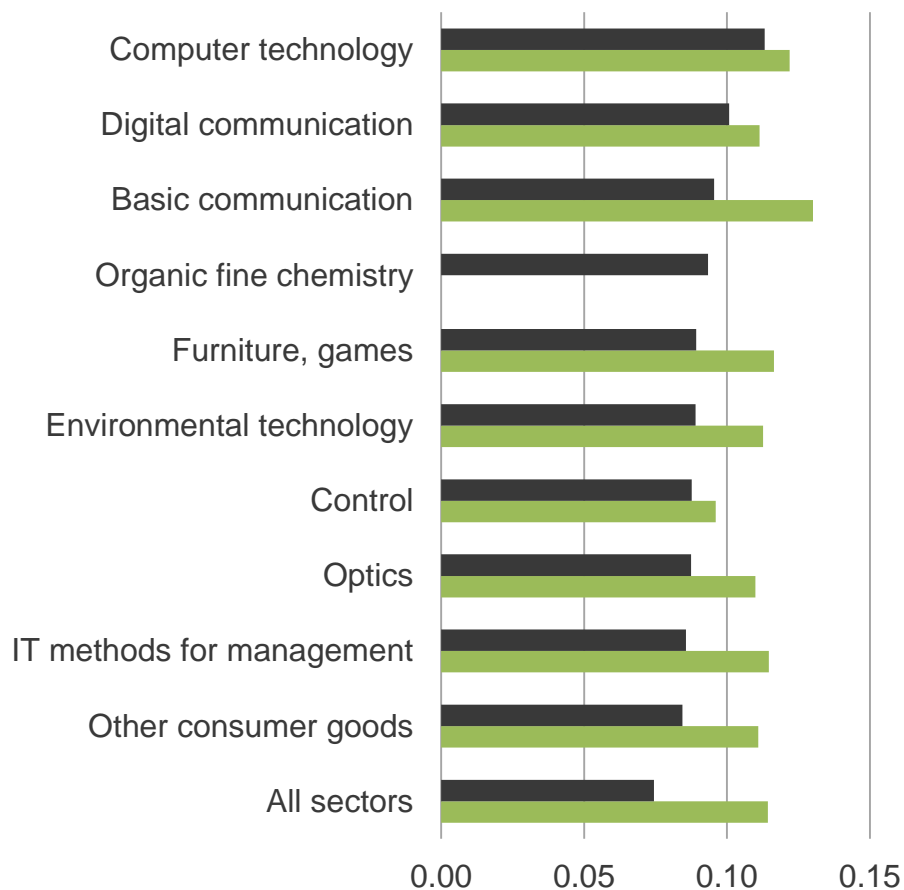


Measuring Patent Quality

Forward citations

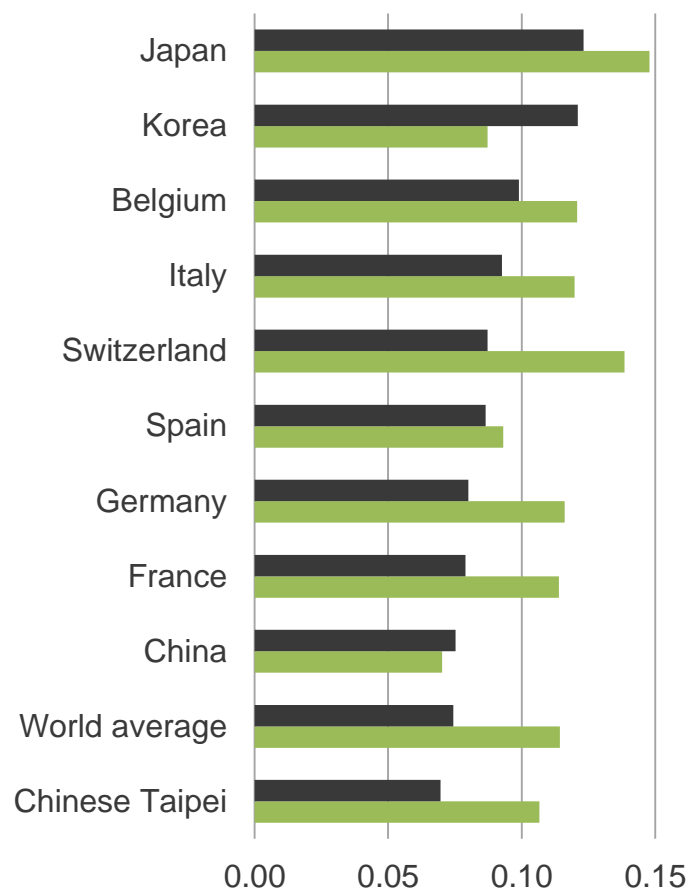
Top 10 technologies, average index

■ 2004 ■ 1994



Top 10 countries, average index

■ 2004 ■ 1994



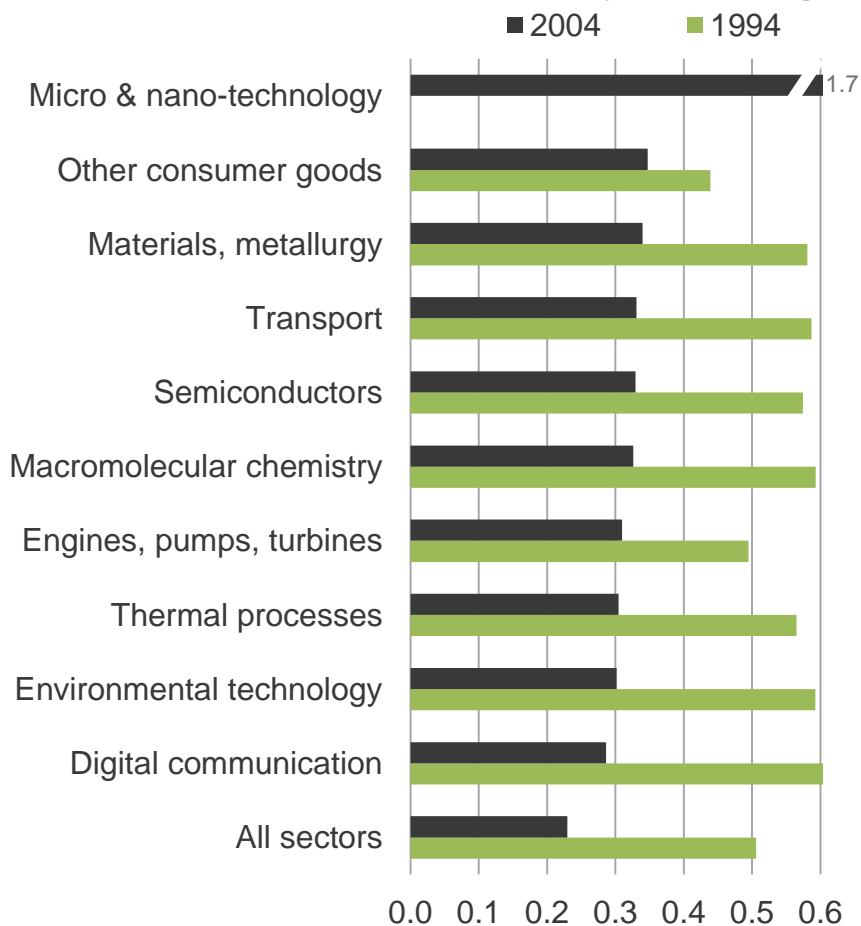
Source: OECD, calculations based on PATSTAT (EPO, April 2012), October 2012



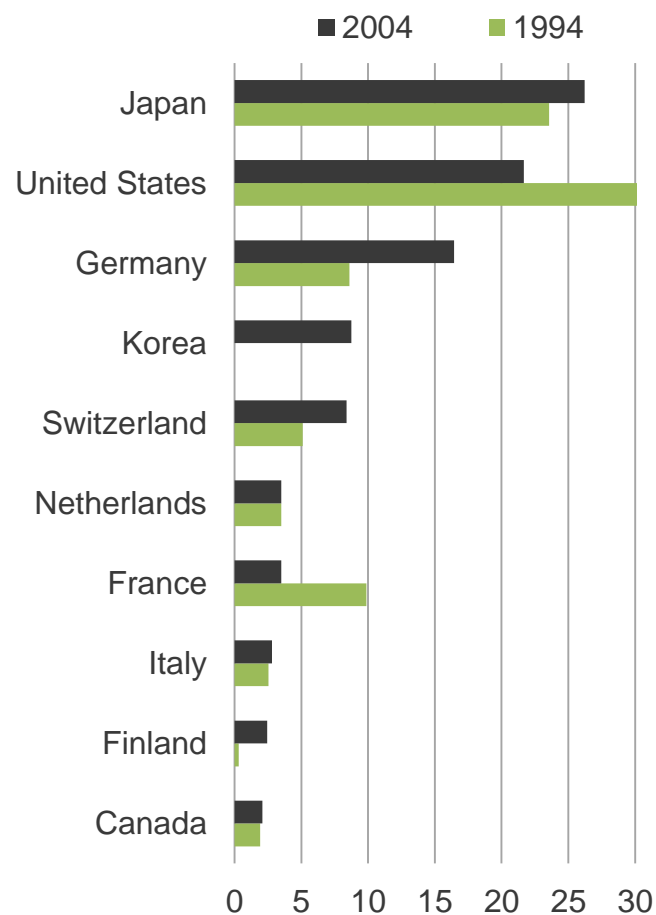
Measuring Patent Quality

Breakthrough patents

Share of breakthrough patents
in total patents, by technologies



Share of countries
in breakthrough patents



Source: OECD, calculations based on PATSTAT (EPO, April 2012), October 2012



Trademark Data

Raw data



Administrative micro-datasets on TM applications at IP Offices worldwide:

- **EU IPO** (ex OHIM): EU Trademarks (ex Community Trademarks)
- **USPTO**
- **JPO** (for internal use only)
- **IP Australia**

Bibliographic details on:

- filing/registration dates,
- applicant,
- types of products protected (Nice class + detailed description)

No worldwide / harmonised dataset available yet

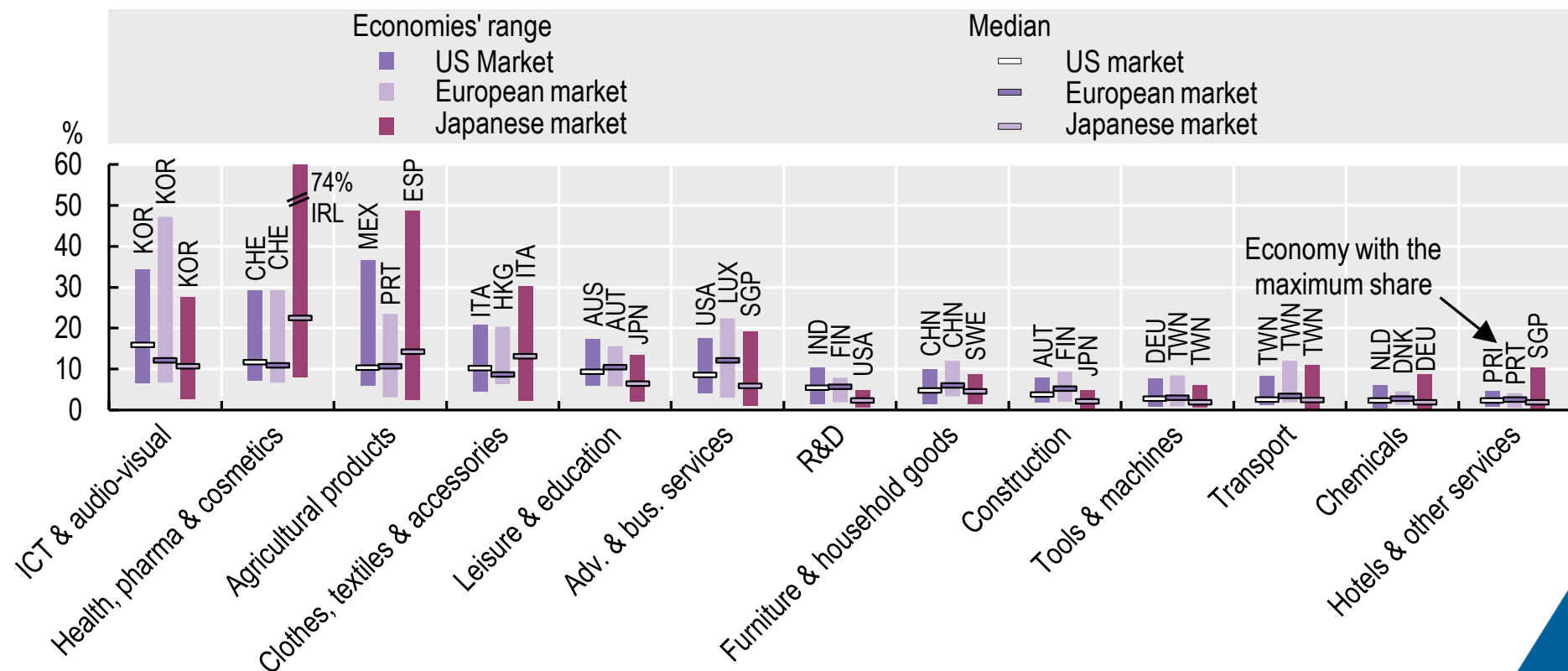


Trademarks

Indicators and Analysis

Trademark specialisation in European, Japanese and US markets, 2012-14

Percentage share of each product field in economies' trademark applications





COMBINING DATASETS



Linking patents to firms

Matching algorithms & HAN dataset

To analyse the links between IP, innovation and firm dynamics, IP records may be matched to company level data.

Patents / TM applicants names were are matched to company names listed in the **ORBIS©** database using a tailored-made algorithm – ***Imalinker*** (Idener Multi Algorithm Linker)

- Matching performed on a country by country basis
- Harmonisation of names with country specific dictionary
- String matching algorithms (token/string metric based)
- High precision (thresholds of high matching scores to minimise the false-positive and false-negative matches)
- Manual controls to adjust the precision.

The result of the matching procedure was used to generate **groupings of patent applicant names** (*HAN – harmonised applicant names*), built on the harmonised name or on the outcomes of the matching.

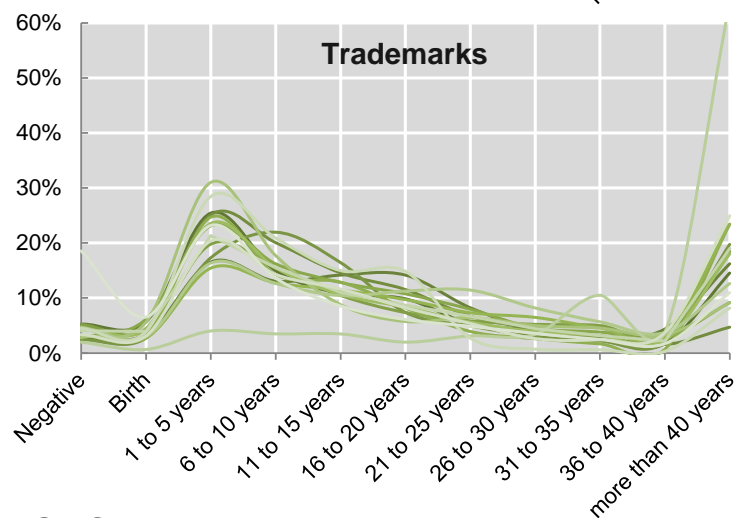
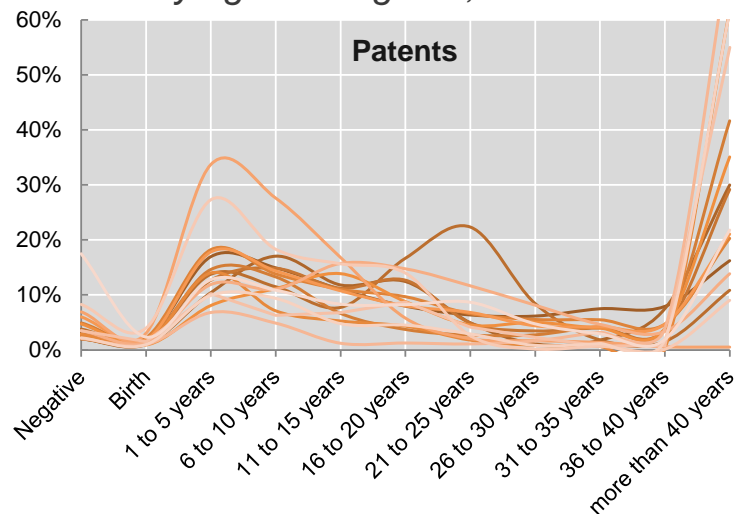
Raw data (OECD HAN Database) available upon request



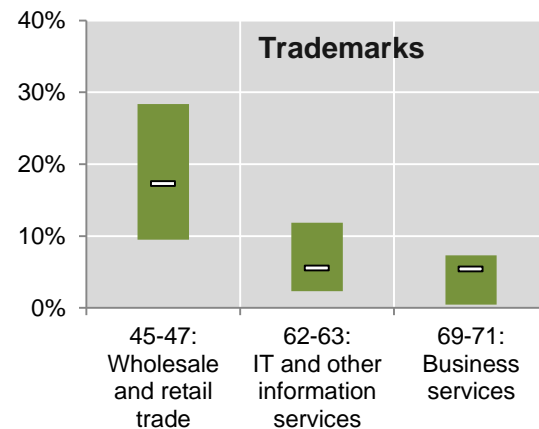
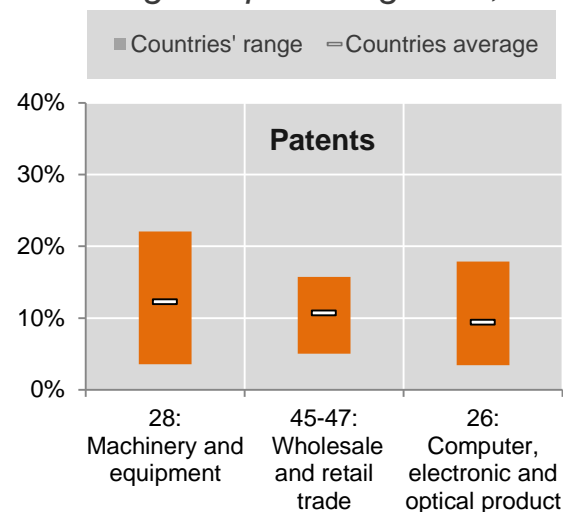
Patenting and trademarking firms

by firm age; by industry

Distribution of patents and trademarks
by age of filing firm, 2000-10



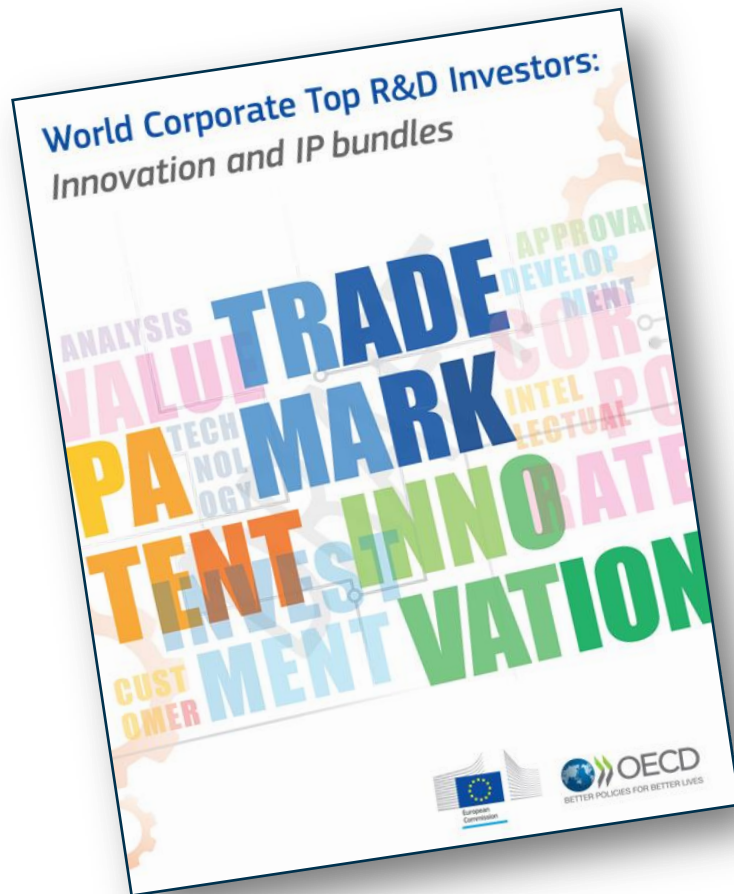
Top 3 industry classes
trademarking and patenting firms, 2000-10





WORLD CORPORATE TOP R&D INVESTORS -

Partnership w/ EC-JRC

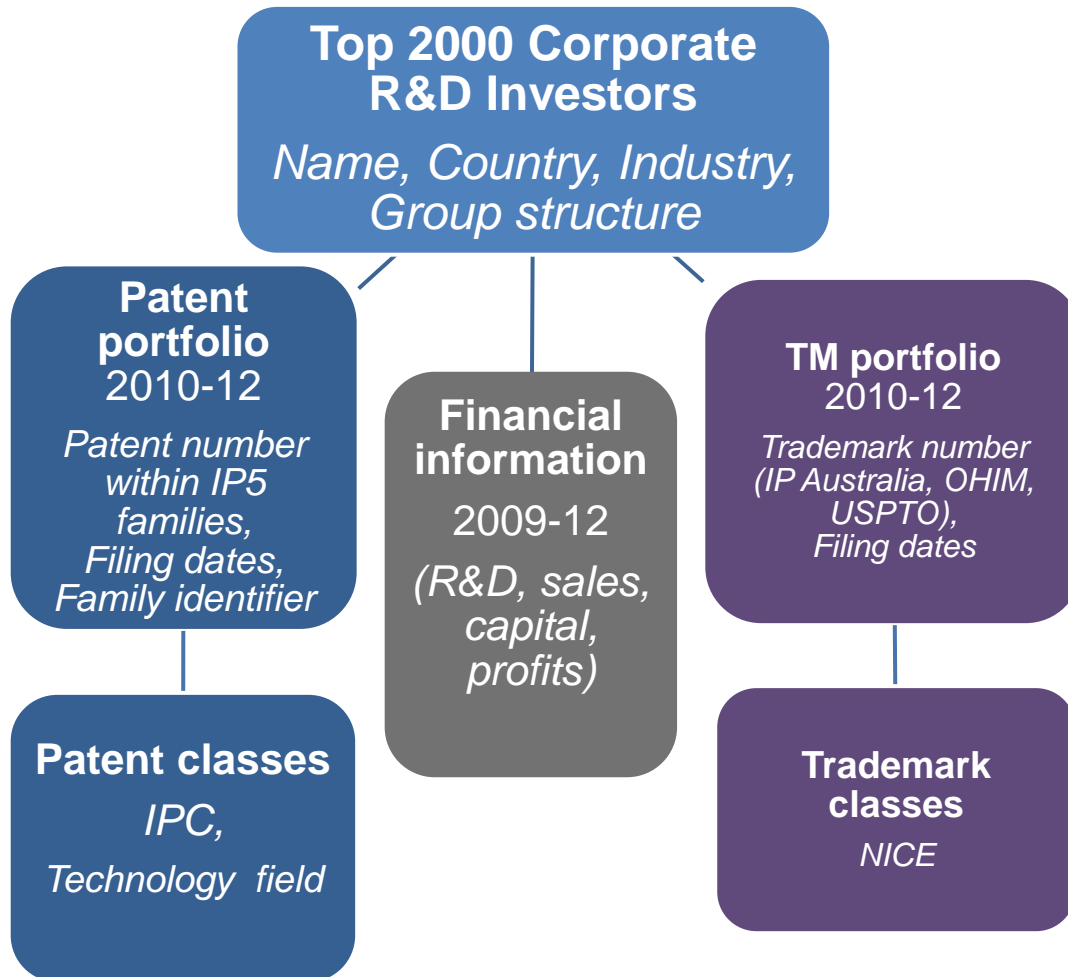


- ❑ **2000 TOP R&D GLOBAL PLAYERS** in 46 economies
- ❑ **+ >500 000 SUBSIDIARIES** in 202 economies (90% of business R&D, 66% of IP5 patents, almost 50% of patent families in electrical engineering)
- ❑ **IP BUNDLE** (IP5 patent families and US, JP, EU trademarks)
- ❑ **OPEN DATA (COR&DIP dataset)**
How to access the data :
 - ❑ Send a request to STI.Microdatalab@oecd.org
 - ❑ Fill in a short on-line form



WORLD CORPORATE TOP R&D INVESTORS -

Partnership w/ EC-JRC



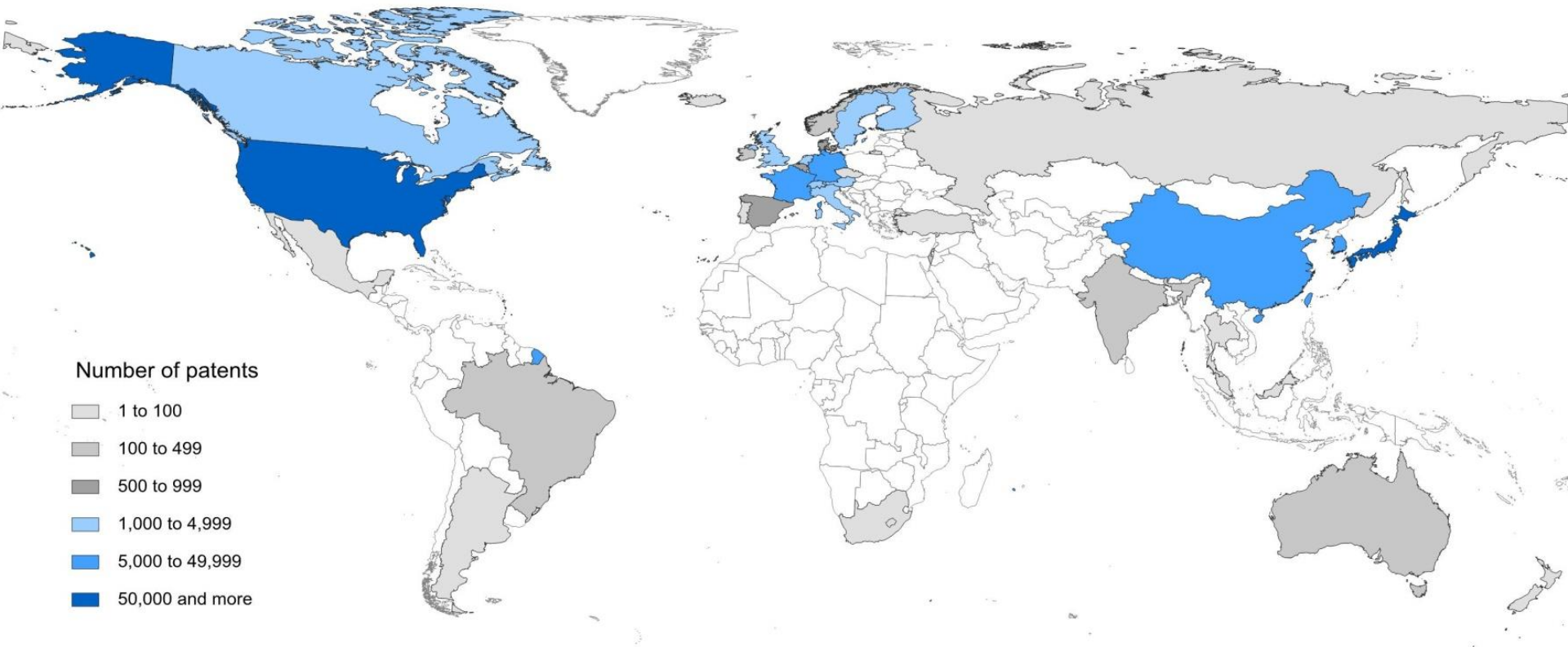
- ☐ **2000 TOP R&D GLOBAL PLAYERS** in 46 economies
- ☐ + **>500 000 SUBSIDIARIES** in 202 economies (90% of business R&D, 66% of IP5 patents, almost 50% of patent families in electrical engineering)
- ☐ **IP BUNDLE** (IP5 patent families and US, JP, EU trademarks)
- ☐ **OPEN DATA (COR&DIP)** How to access the data :
 - ☐ Send a request to STI.Microdatalab@oecd.org
 - ☐ Fill in a short on-line form

Watch out for 2017 edition!



World top R&D investors

Location of patenting firms, 2012



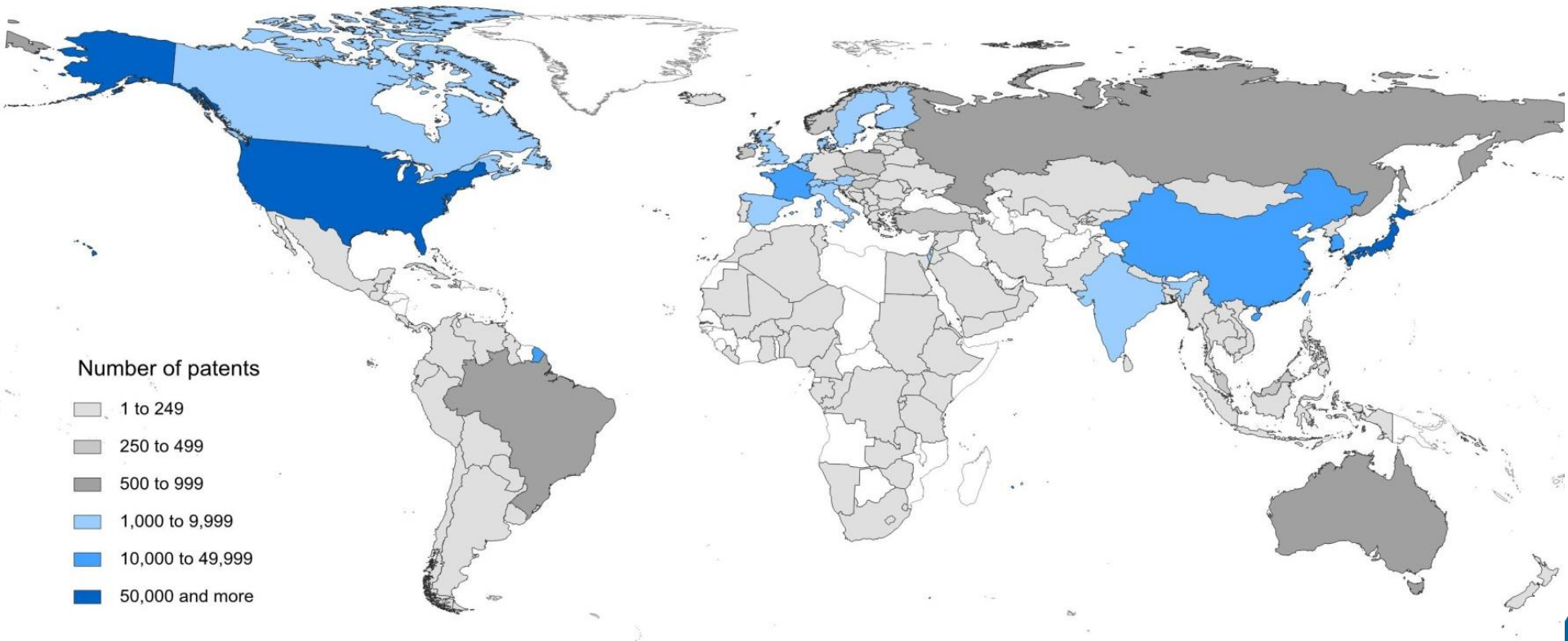
Source: IPTS-OECD, World Corporate Top R&D Investors: Innovation and IP bundles, 2015.

Map source: ARCTIQUE© - All rights reserved.



World top R&D investors

Location of inventors, 2012



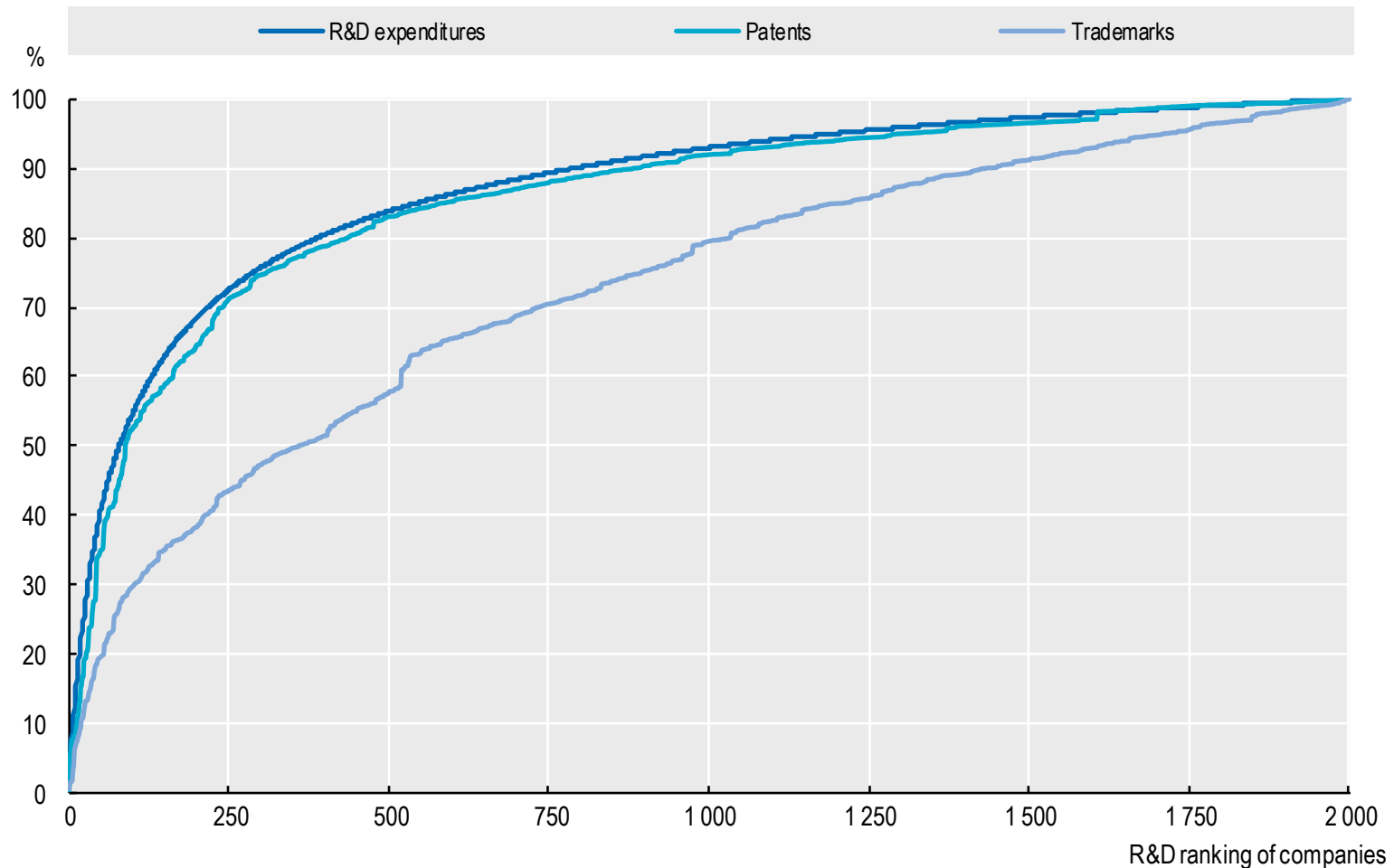
Source: IPTS-OECD, World Corporate Top R&D Investors: Innovation and IP bundles, 2015.

Map source: ARCTIQUE© - All rights reserved.



Top 250 R&D companies account for over 60% of global business R&D, 2012

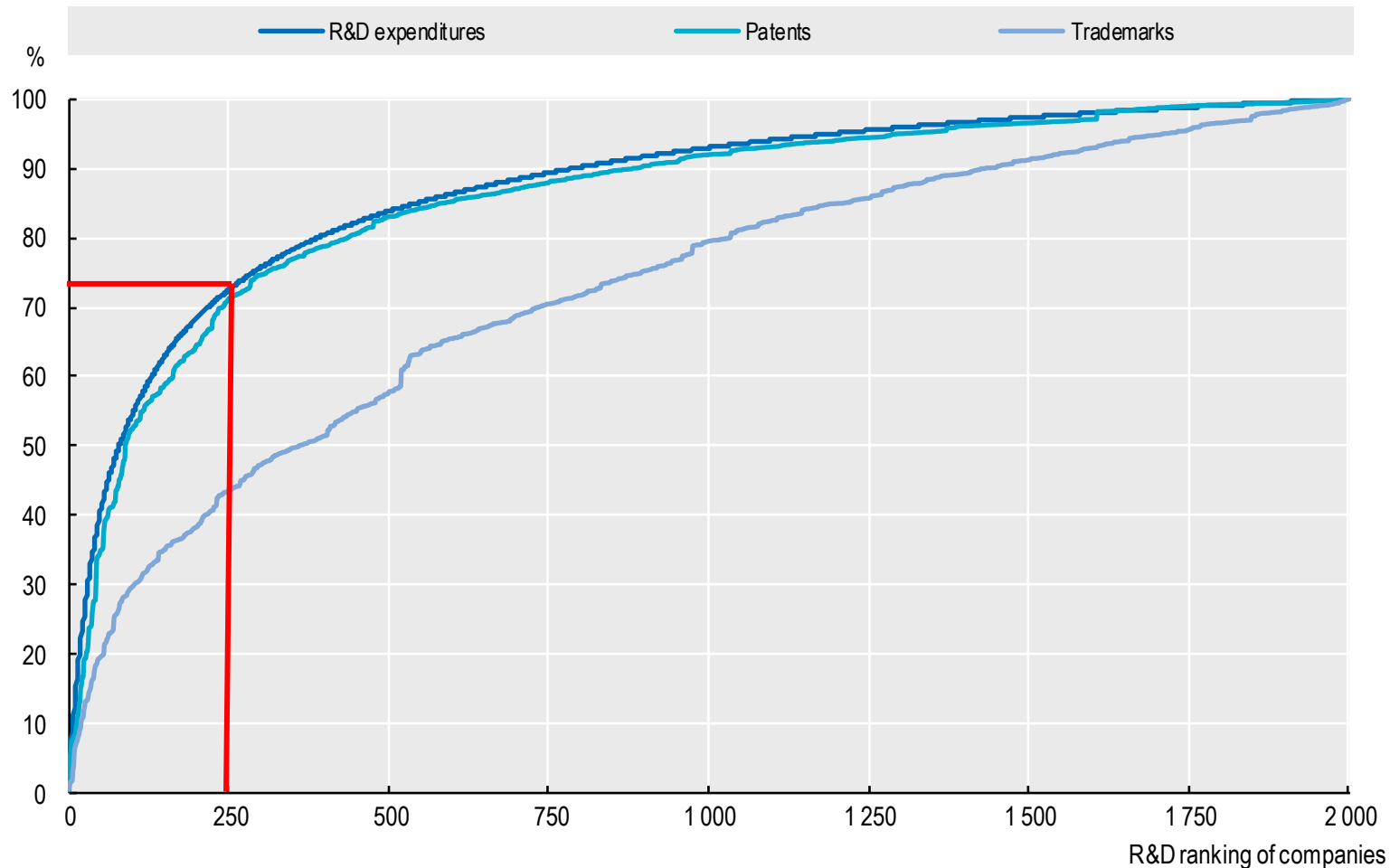
Cumulative percentage shares within the top 2000 R&D companies





Top 250 R&D companies account for over 60% of global business R&D, 2012

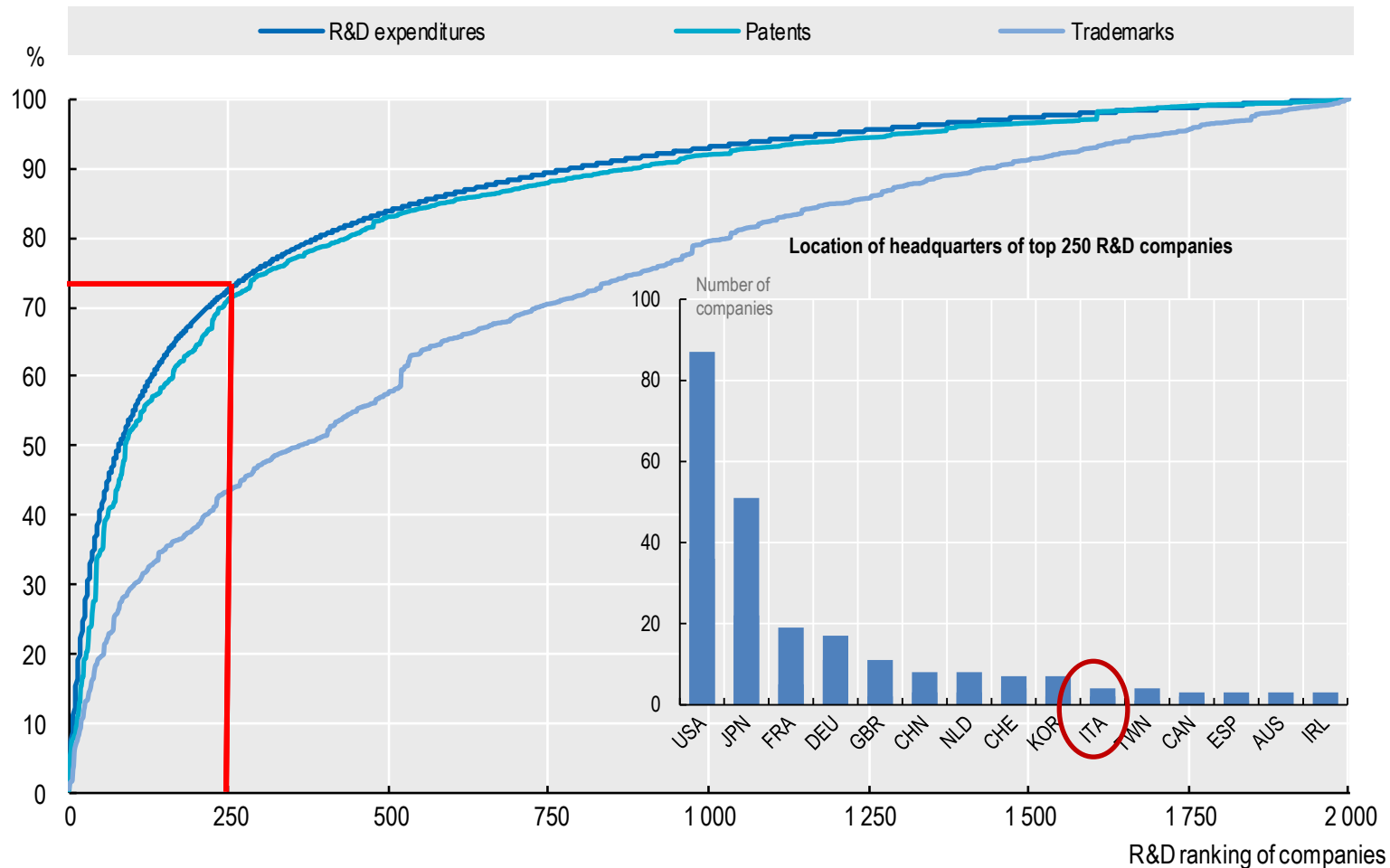
Cumulative percentage shares within the top 2000 R&D companies





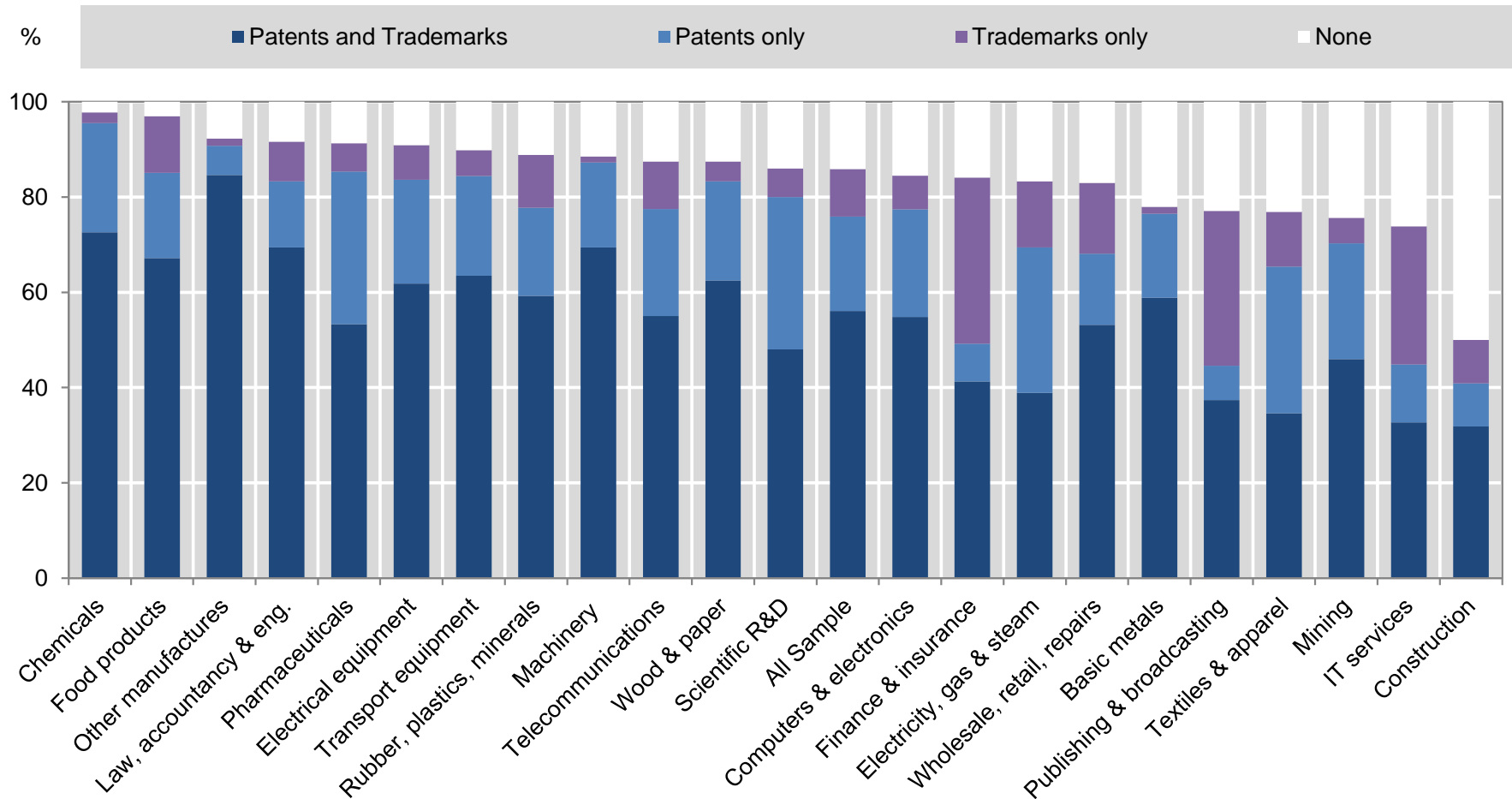
Top 250 R&D companies account for over 60% of global business R&D, 2012

Cumulative percentage shares within the top 2000 R&D companies





Top R&D investors with TM and patents, by industry, EPO and OHIM, 2010-12



Source: IPTS-OECD, World Corporate Top R&D Investors: Innovation and IP bundles, 2015.



Technology advantage

by headquarter location, 2010-12

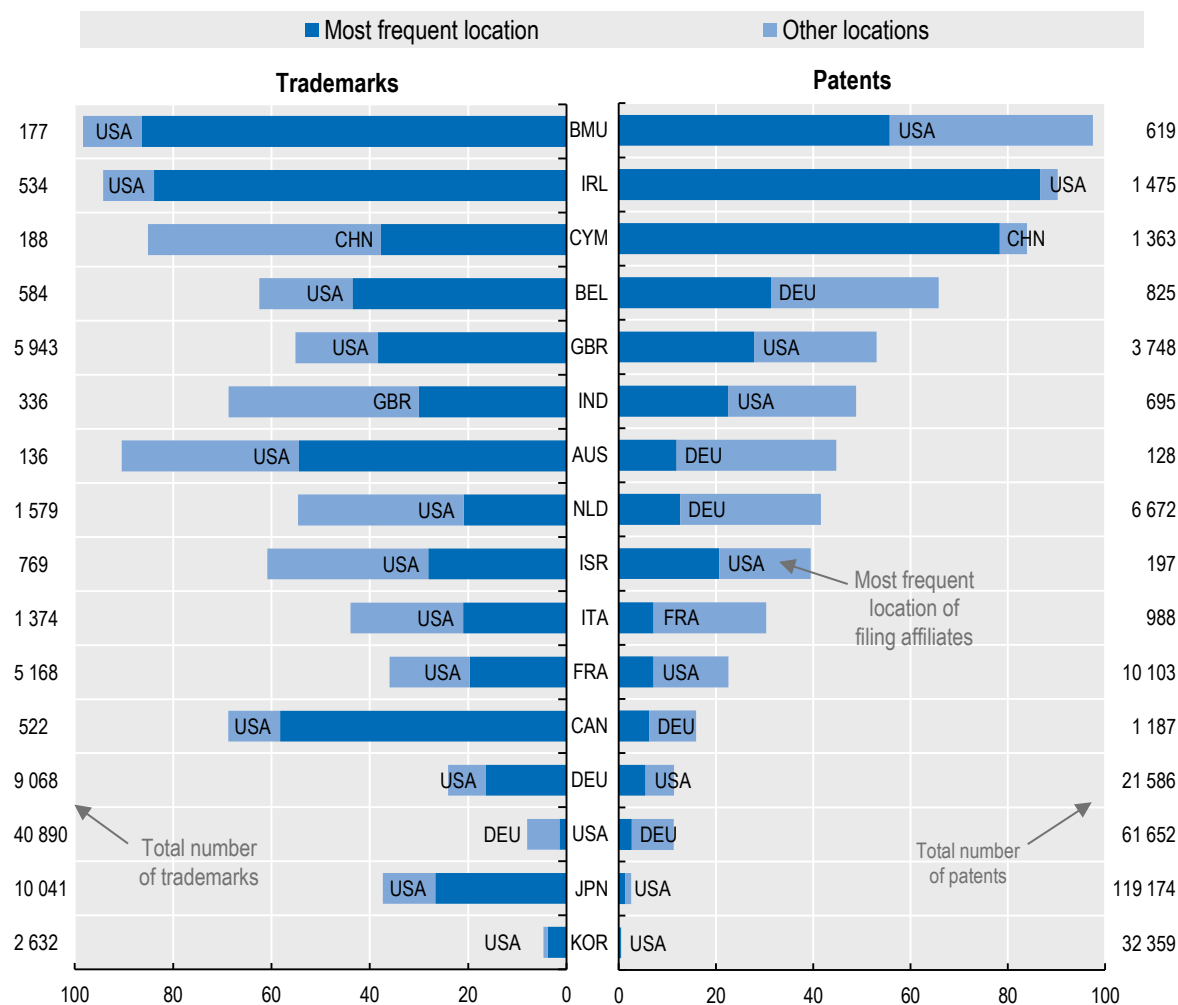
Field of Technology	Europe	United States	Japan	Korea	China	RoW
Electrical machinery	1.0	0.7	1.1	1.3	0.5	1.1
Audio-visual tech.	0.4	0.5	1.2	1.6	0.6	2.1
Telecommunications	0.7	0.7	1.0	1.4	3.1	1.3
Digital communication	1.1	1.1	0.6	1.3	8.0	1.2
Basic communication	0.8	1.0	1.0	1.0	1.1	1.7
Computer technology	0.5	1.3	0.8	1.4	1.4	1.8
IT methods	0.8	1.8	0.7	1.0	0.6	1.2
Semiconductors	0.4	0.7	1.1	2.0	0.1	1.5
Optics	0.3	0.4	1.6	1.1	0.2	1.0
Micro- and nano-tech.	1.2	1.0	0.7	1.3	0.0	1.7
Measurement	1.4	1.1	0.9	0.5	0.3	0.8
Bio materials	1.6	1.6	0.7	0.6	0.0	0.1
Control	1.7	1.9	0.4	0.1	0.7	1.3
Medical technology	1.5	1.6	0.9	0.3	0.0	0.2
Organic chemistry	2.0	1.4	0.6	0.3	0.5	0.3
Biotechnology	1.8	1.6	0.6	0.6	0.1	0.2
Pharmaceuticals	2.0	1.8	0.5	0.2	0.1	0.6

Source: IPTS-OECD, World Corporate Top R&D Investors: Innovation and IP bundles, 2015.



IP filings by foreign affiliates of top R&D corporations

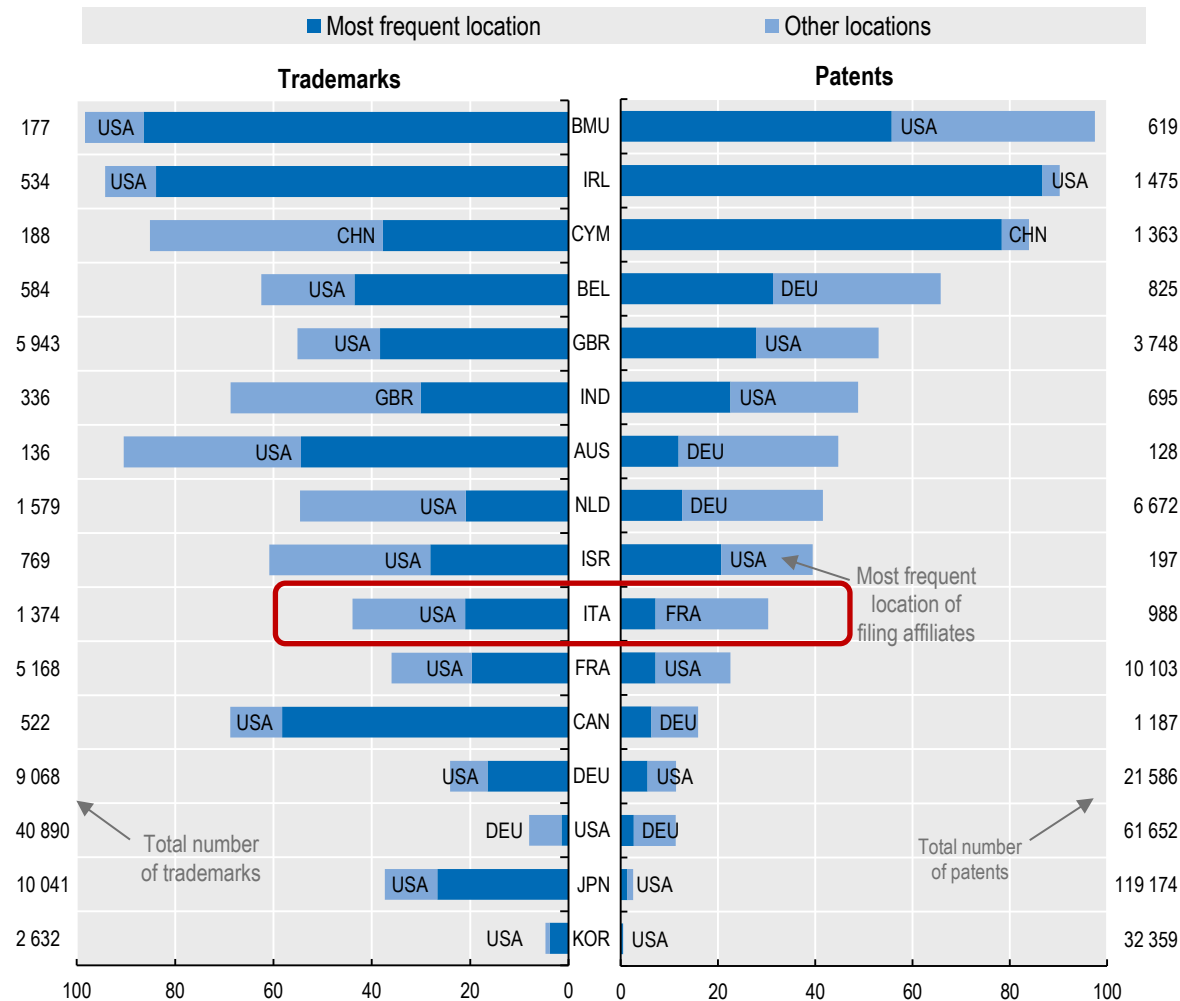
By location of the headquarters, 2010-12





IP filings by foreign affiliates of top R&D corporations

By location of the headquarters, 2010-12





EMERGING TECHNOLOGIES



Patents and emerging technologies

OECD definitions

- OECD analysis focusing on selected technology areas
- Definitions for patents in **ICT**, **Biotech**, **Environment** related technologies, based on the IPC or CPC
- **New definition for ICT** to reflect recent evolution in the field
- Ongoing **revision of Biotech** patents taxonomy
- OECD ENV Directorate identified sets of **environment-related** technologies
- **New text mining** approach to identify emerging technologies

Dernis, Squicciarini & de Pinho (2016)

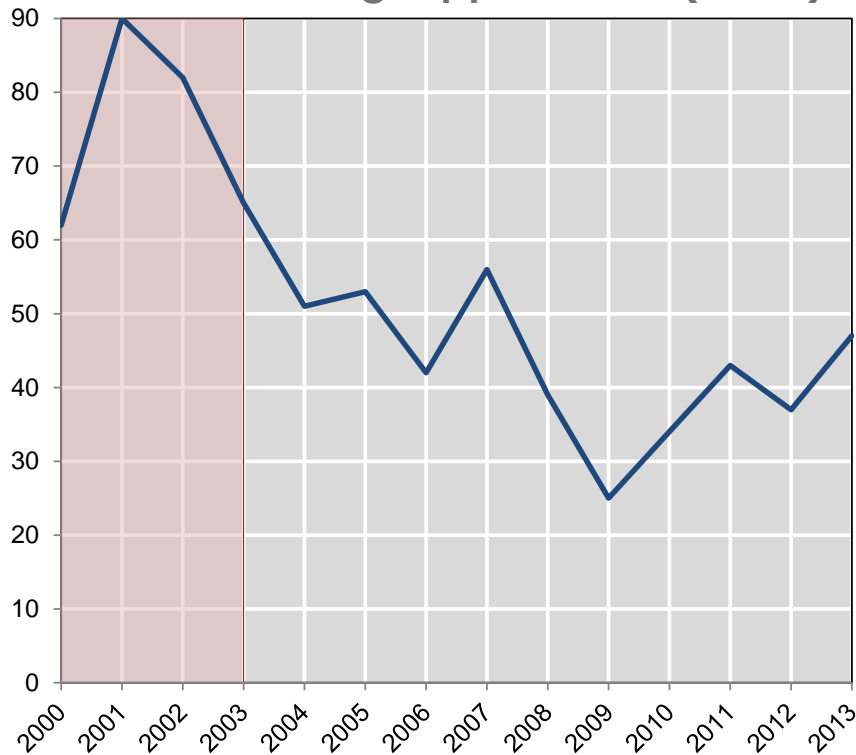
Detecting the emergence of technologies and the evolution and co-development trajectories in science,
Journal of Technology Transfer





Detecting the emergence and decline of technologies

Ticket-issuing apparatuses (G07B)

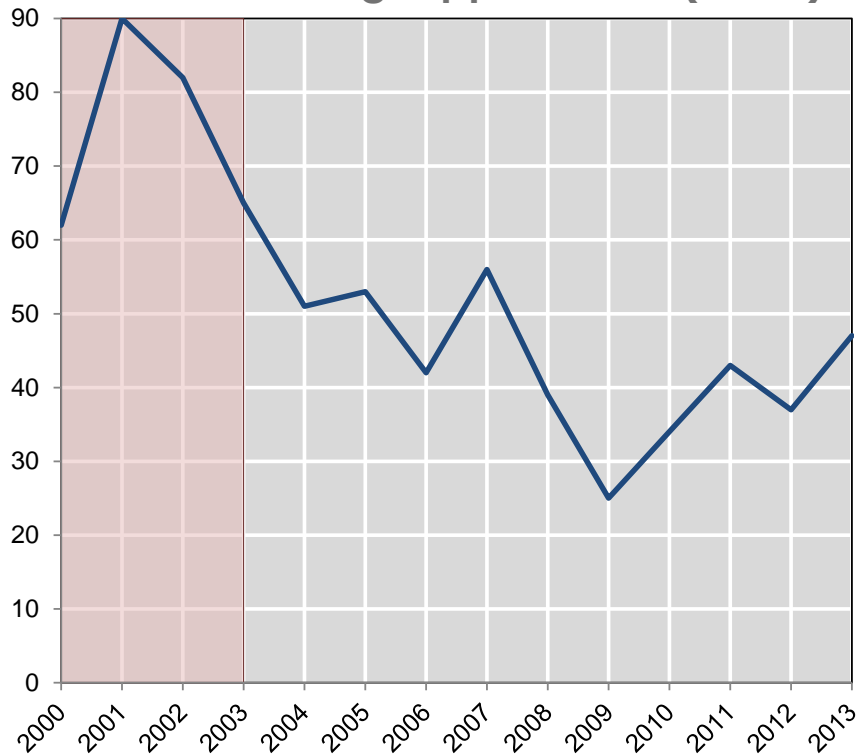


Source: [Dernis, Squicciarini and De Pinho](#), Journal of Technology Transfer (2016)

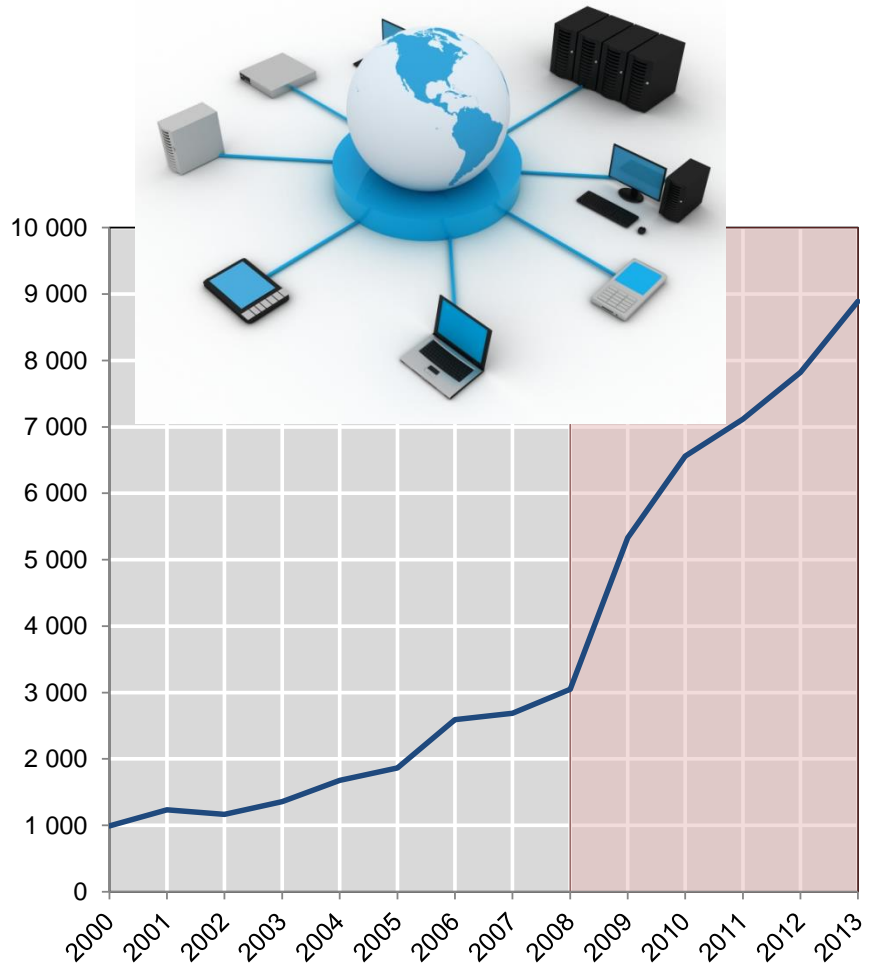


Detecting the emergence and decline of technologies

Ticket-issuing apparatuses (G07B)



Wireless communication networks (H04W)

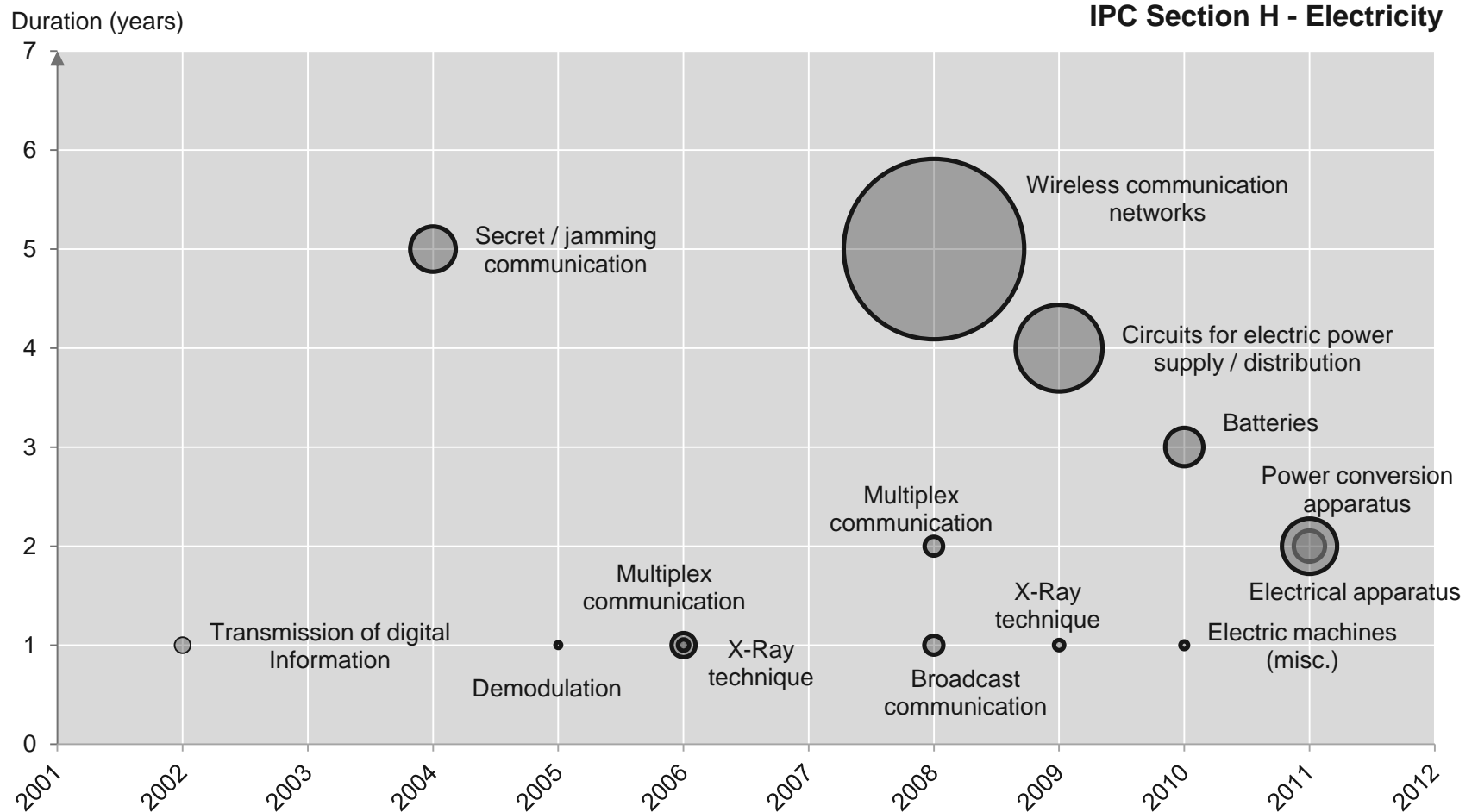


Source: [Dernis, Squicciarini and De Pinho](#), Journal of Technology Transfer (2016)



Burst intensity and duration

by IPC subclass and first year of burst



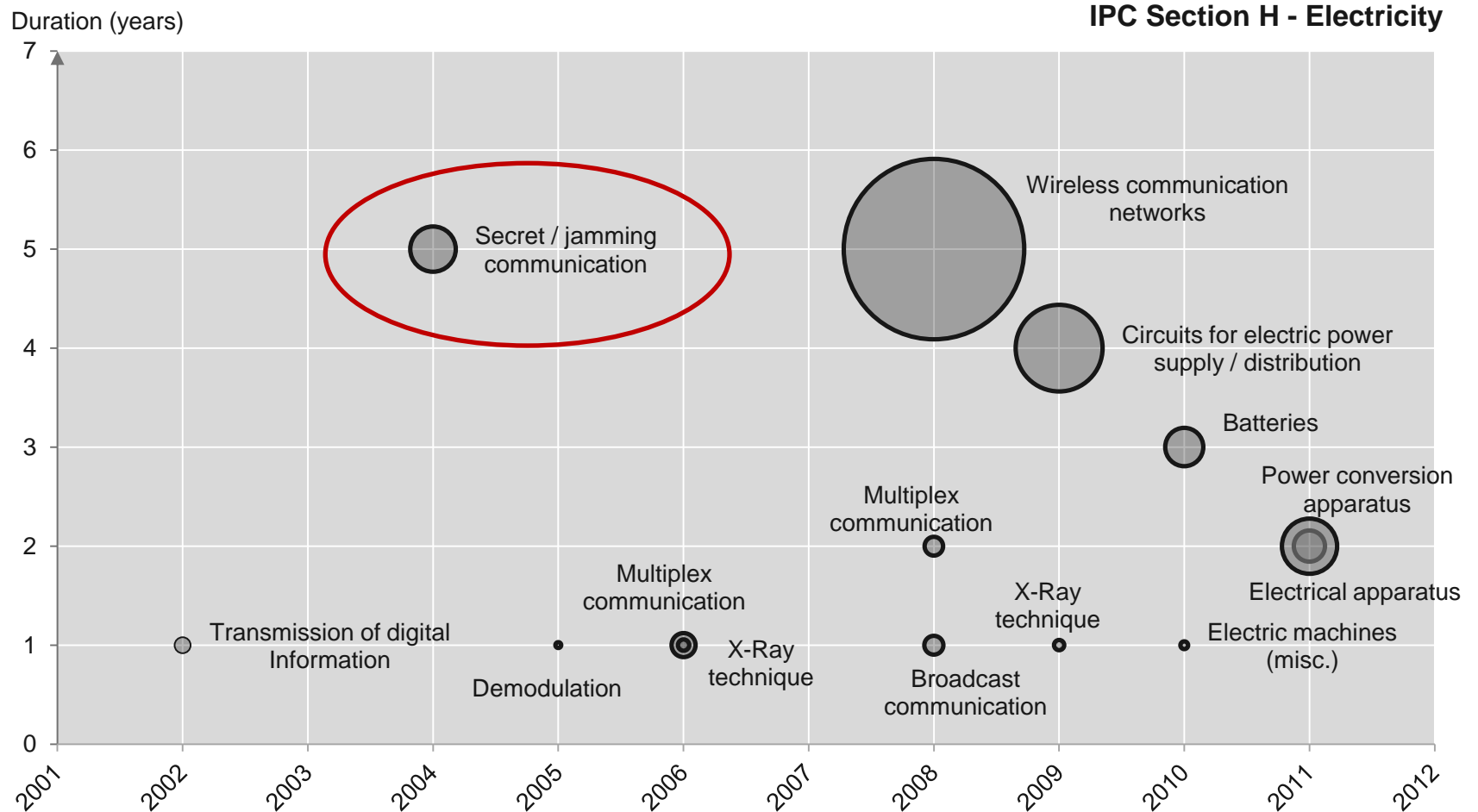
Note: Burst intensities observed in IP5 Patent Families.

Source: [Dernis, Squicciarini and De Pinho](#), Journal of Technology Transfer (2016)



Burst intensity and duration

by IPC subclass and first year of burst



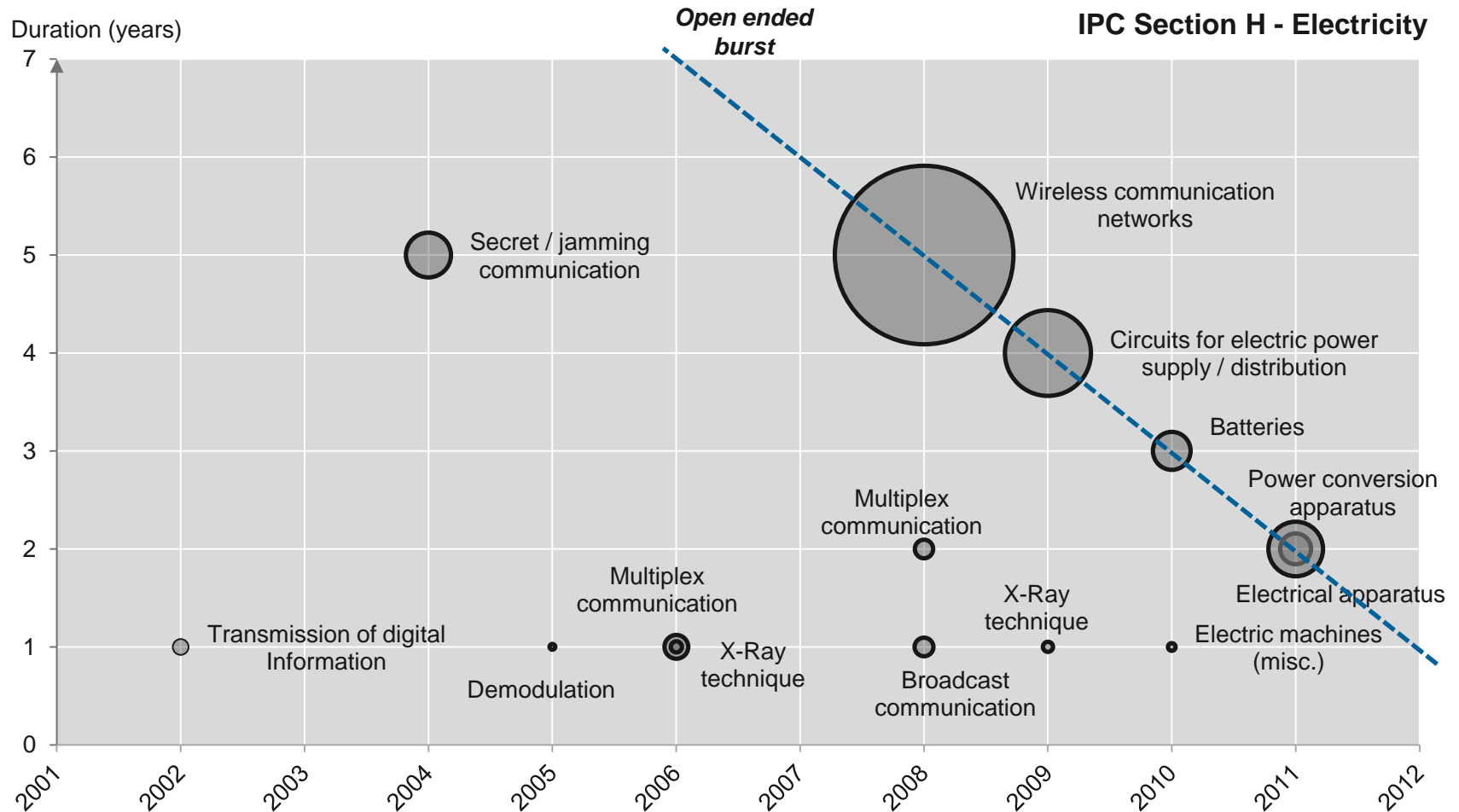
Note: Burst intensities observed in IP5 Patent Families.

Source: [Dernis, Squicciarini and De Pinho](#), Journal of Technology Transfer (2016)



Burst intensity and duration

by IPC subclass and first year of burst



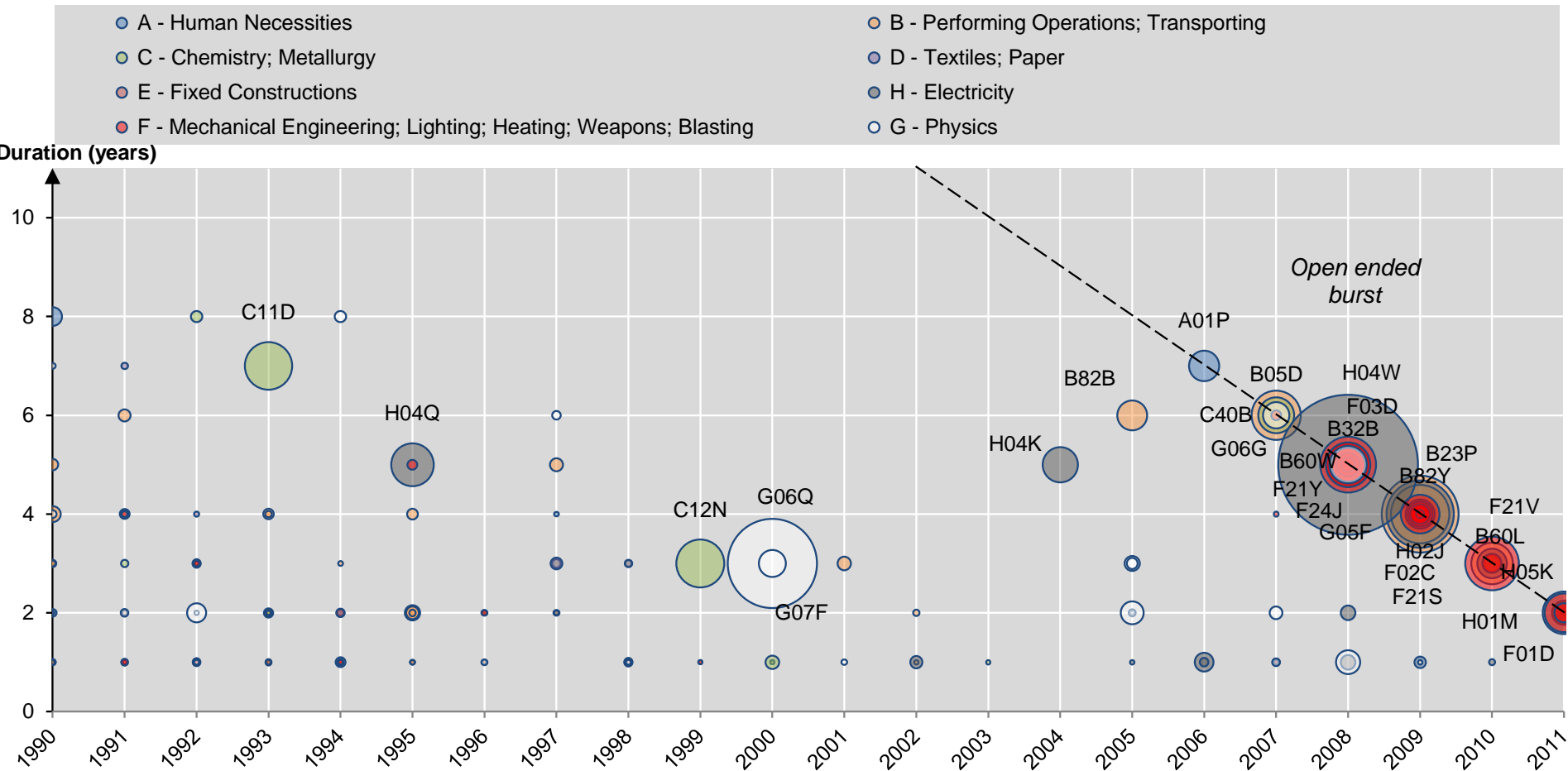
Note: Burst intensities observed in IP5 Patent Families.

Source: [Dernis, Squicciarini and De Pinho](#), Journal of Technology Transfer (2016)



Burst intensity and duration

by IPC subclass and first year of burst



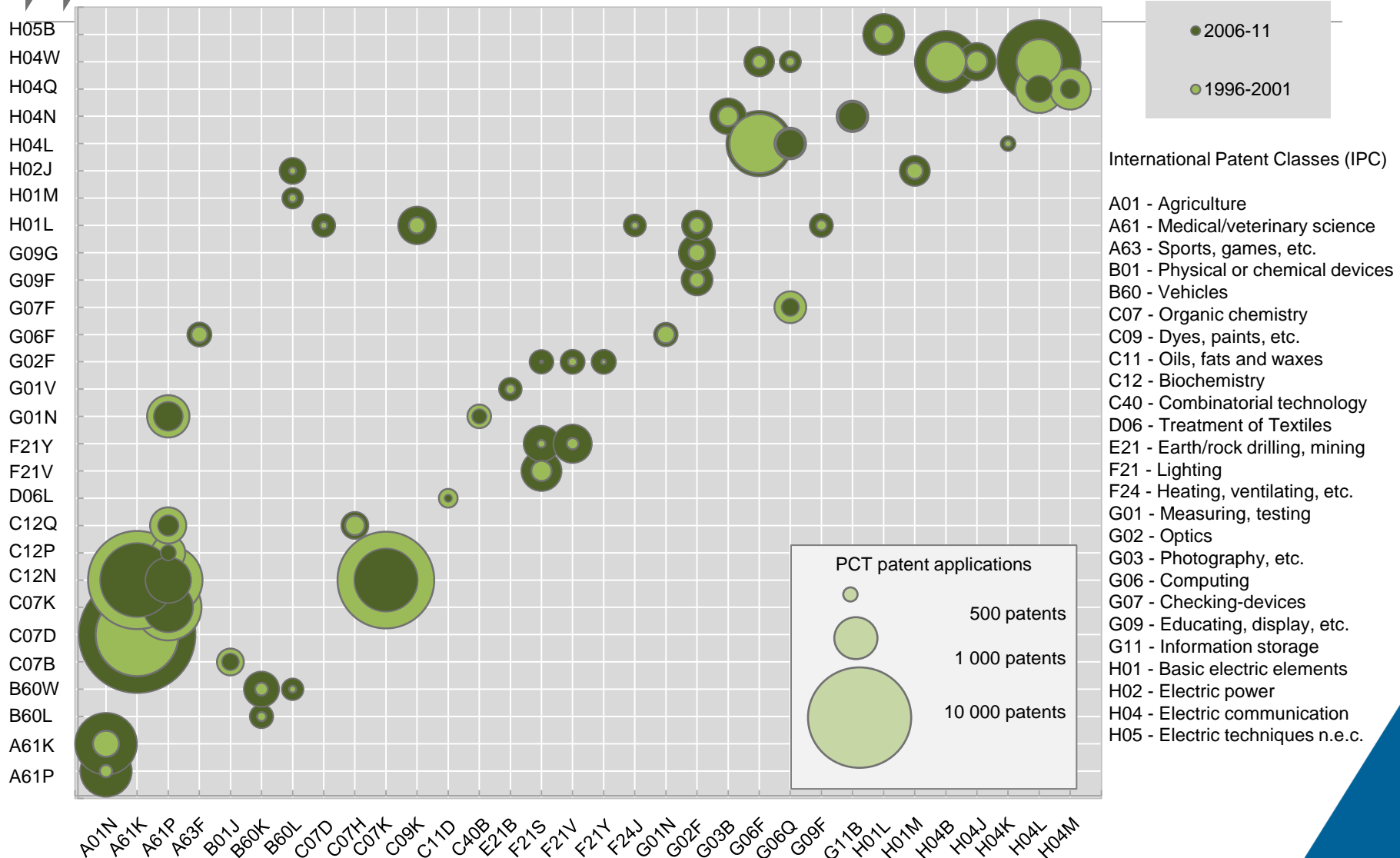
Note: Burst intensities observed in IP5 Patent Families

Source: Author's calculations based on EPO PATSTAT, Spring 2014, October 2014.



Co-development of technologies

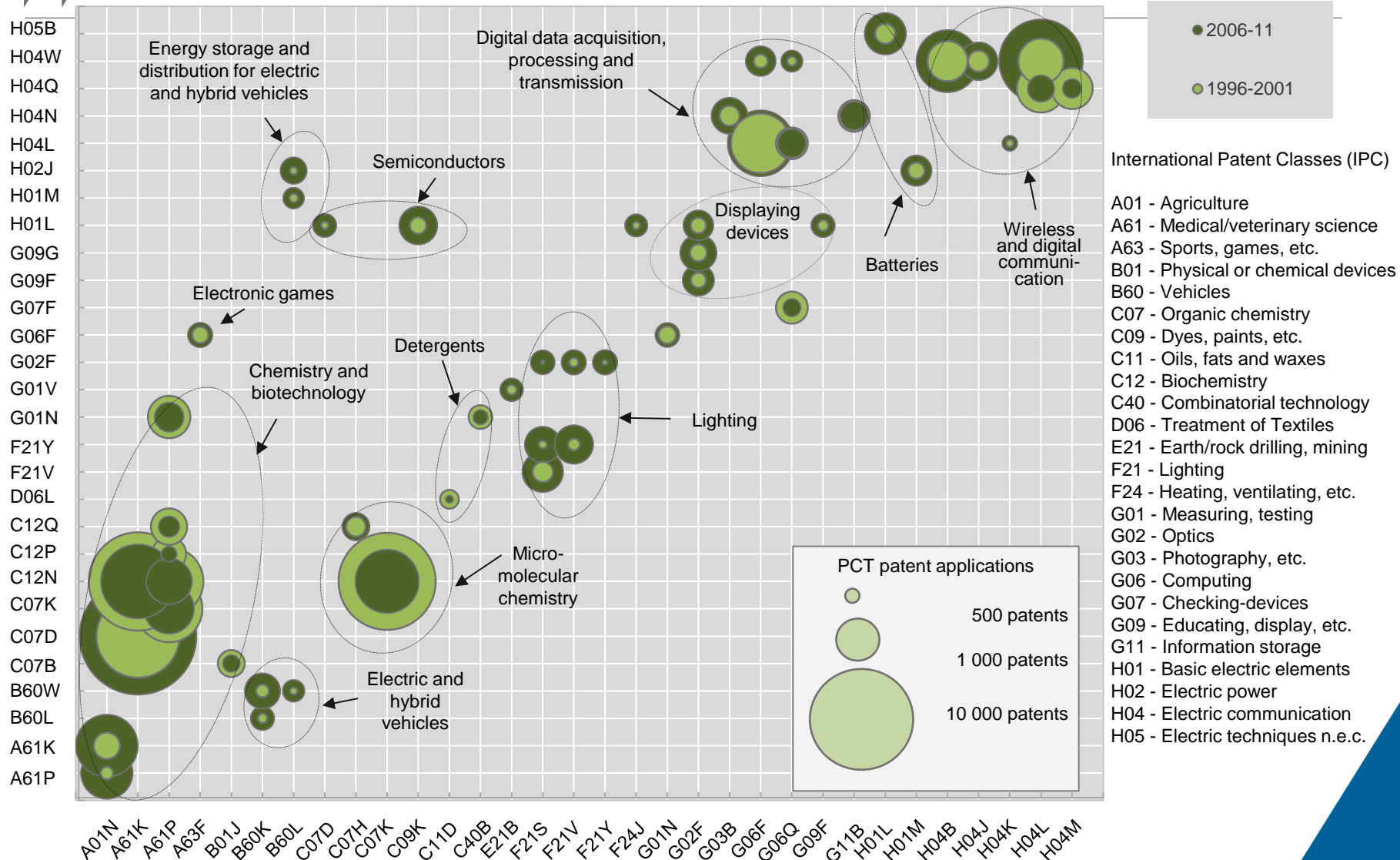
New fields arising from cross-fertilisation





Co-development of technologies

New fields arising from cross-fertilisation

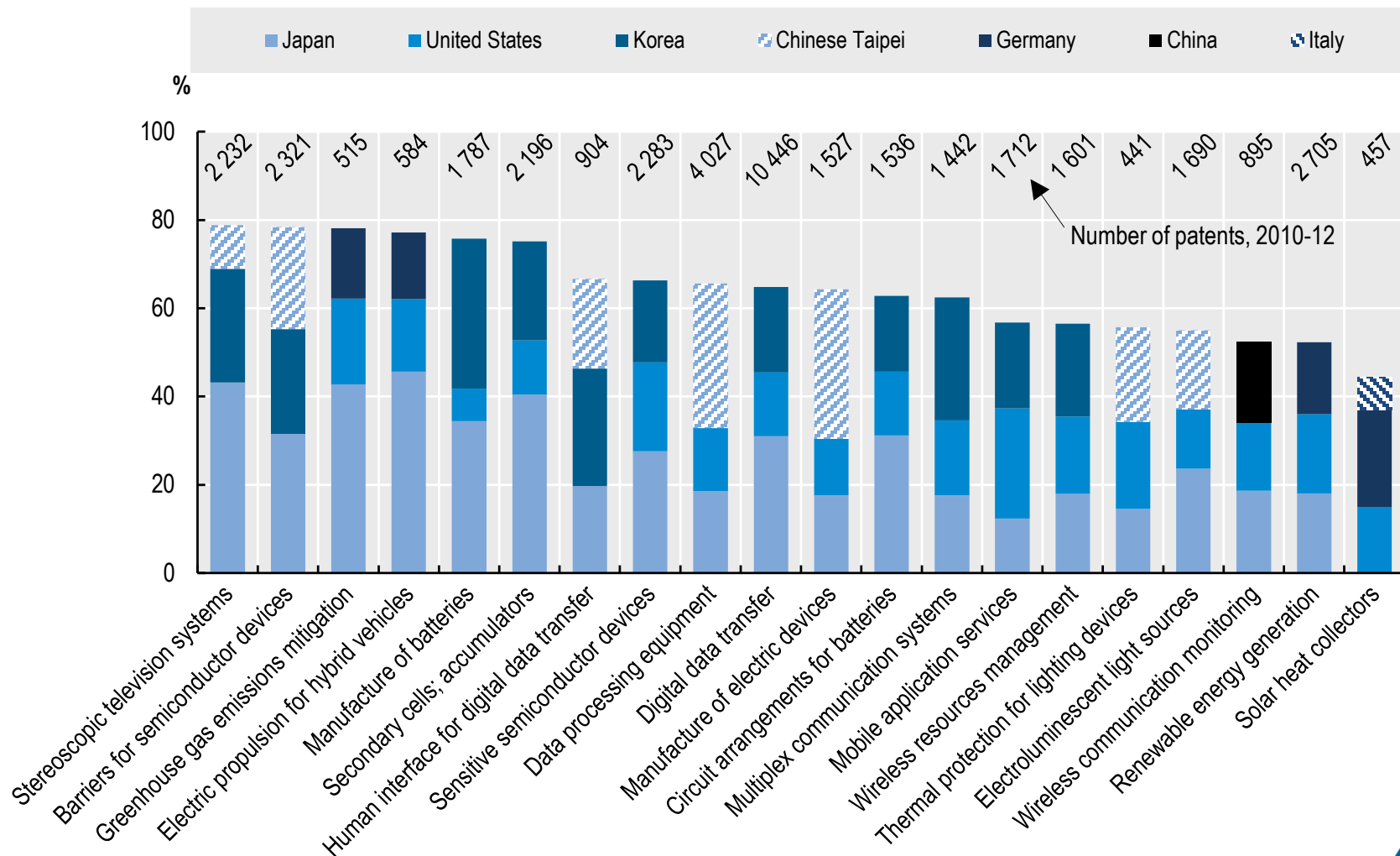




Top players in emerging technologies

2010-12

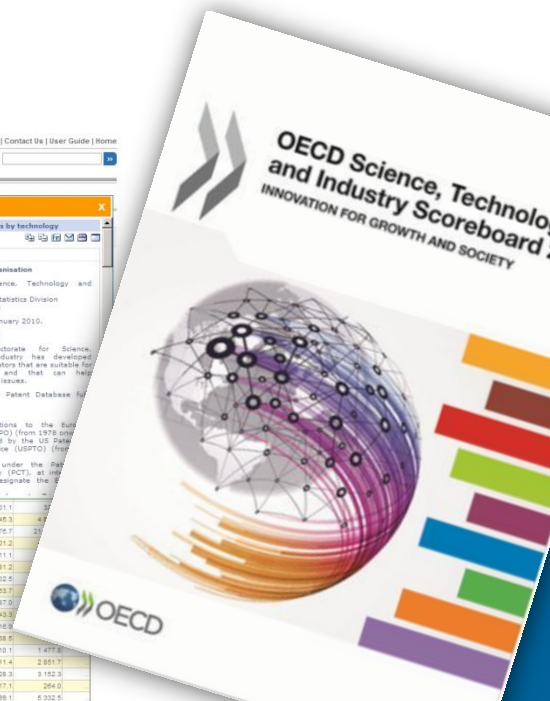
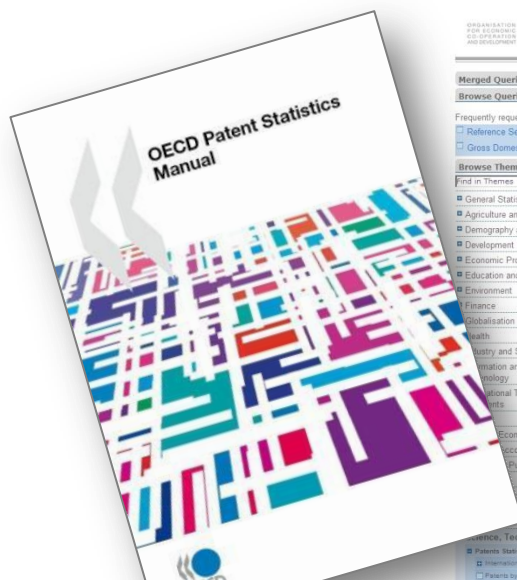
Top three economies' patents in top 20 technologies bursting from 2005



OECD STI Microdata lab

Work in progress

- Analysis of new emerging fields
- Further develop indicators on the IP bundle (TM / Design)
- Concordance IP-related classifications vs industries
- Launching a survey on the IP systems
- Raw datasets (patents) available upon request
sti.microdatalab@oecd.org
- Further information: oe.cd/ipstats





Grazie!

Contact: mariagrazia.squicciarini@oecd.org
www.researchgate.net/profile/Mariagrazia_Squicciarini

Per maggiori informazioni:

Website: www.oecd.org/sti
oe.cd/ipstats

Twitter: [@OECDInnovation](https://twitter.com/OECDInnovation)

Newsletter: www.oecd.org/sti/news.htm