

Issues in measuring the degree of technological specialisation with patent data

N. van Zeebroeck, B. van Pottelsberghe de la Potterie and W. Han

This paper analyses several issues that arise when measuring technological specialisation with patent data. Three starting choices are required regarding the data source, the statistical measure and the sectoral aggregation level. We show that the measure is highly sensitive to the data source and to the level of sectoral aggregation. The statistical analysis further suggests that the most stable and reliable measures of technological specialization are obtained with patents applied at the EPO, with Gini or C20 as statistical measure and the 4-digits aggregation level of the IPC classification system.

JEL Classifications: L16; O3; O57

Keywords: Technological specialisation, patent data, patents statistics

CEB Working Paper N° 05/016
April 2005

Issues in measuring the degree of technological specialisation with patent data

Nicolas van Zeebroeck ^σ, Bruno van Pottelsberghe de la Potterie ^α, Wook Han [∇]

Revised version: April 2005

Published in Scientometrics, Vol. 66, No. 3 (February 2006), pp. 481-492

Suggested Citation

van Zeebroeck, Nicolas, van Pottelsberghe de la Potterie, Bruno and Han, Wook, "Issues in measuring the degree of technological specialisation with patent data" (February 2006). Scientometrics, Vol. 66, No. 3, pp. 481-492. Available at: <http://www.akademai.com/link.asp?id=2w387205g231n128>.

^σ ULB – Solvay Business School, Centre Emile Bernheim (CEB) – Av. Roosevelt 21 | 1050 Brussels (Belgium) – nicolas.van.zeebroeck@ulb.ac.be

^α ULB – Solvay Business School, Centre Emile Bernheim (CEB), DULBEA and CEPR – Av. Roosevelt 21 | 1050 Brussels (Belgium) – bruno.vanpottelsberghe@ulb.ac.be

[∇] ULB – Solvay Business School

Abstract

This paper analyses several issues that arise when measuring technological specialisation with patent data. Three starting choices are required regarding the data source, the statistical measure and the sectoral aggregation level. We show that the measure is highly sensitive to the data source and to the level of sectoral aggregation. The statistical analysis further suggests that the most stable and reliable measures of technological specialization are obtained with patents applied at the EPO, with Gini or C20 as statistical measure and the 4-digits aggregation level of the IPC classification system.

Author Keywords: Technological specialisation, patent data, patents statistics

JEL classification codes: L16; O3; O57

1. Introduction

Patents and patent statistics have been intensely used by scholars attempting to measure technological change and the sources of economic growth. Patent data appeared as an important input to scholars attempting to measure technological specialisation, change or innovation performances. Easily and widely available over long time periods and in many countries, they contain detailed information (IPC classes, inventors, etc.) and are by definition related to inventiveness (see Griliches (1990) for an in-depth analysis of patent data as indicator of innovation performances).¹

The technological specialisation of a country can be measured through the distribution of its patents over various technological areas. The more concentrated the patents are in certain technological areas, the more the country is said to be “technologically specialised”. To the best of our knowledge, the first uses of patent data to measure technological change have been performed by Comanor and Scherer (1969) and by Schmookler (1972). Since then, patent data has been used by an increasing number of authors and institutions to measure technological change and specialization, such as Archibugi and Pianta (1992), Braun et al. (1995), the OECD (1994) and the European Commission (1997 and 2003).²

This paper focuses on the most important choices that must be made when using patent data to measure the technological specialisation. In this respect, three dimensions have to be taken into account: the source of patent data, the statistical index used to measure the degree of technological specialisation and the patent classification.

In section 2 we present the theoretical background and the three dimensions of the measure. In section 3 we describe our data sets and our methodology to compare the measures of technological specialisation obtained from different choices along the dimensions. Section 4 presents the statistical results of our comparisons and in Section 5 we conclude with some recommendations on what appears to be the most optimal choices, if any.

2. Three starting choices

A first methodological choice concerns the data source. There are many different sources of patent data available around the world: the United States Patent and Trademark Office (USPTO), the Japanese Patent Office, the European Patent Office (EPO), Triadic Patent Families and the World Intellectual Property Office (WIPO), not to mention all the national patent offices.

The present paper relies on patent data from three different sources: USPTO data containing all national and foreign patents granted in the US, EPO data containing all patents filed at the European Patent Office and Triadic Patent Families gathered by the OECD. As mentioned in EC’ Second Report on S&T Indicators (1997), Triad Patent Families, which consider the patents that have been applied simultaneously at the USPTO, the EPO and the Japanese Patent Office (JPO), derive from the geographical limitation of the protection granted from a domestic patent, and hence from the possible necessity to duplicate applications at overseas patent offices in order to obtain protection in foreign markets. “*The selection of patent data from only one patent office, therefore, does not always yield an indicator that is representative of the world output of inventions.*”³ It is worth noticing that EPO patents

¹ There are, however, several drawbacks to using patent data in order to build economic indicators: they differ greatly in their technical and economic significance; not all inventions are patented, nor all are patentable; other existing methods in appropriating an innovation such as industrial secrecy may be preferred; the propensity to patent may change substantially over time and across countries not to mention among technological sectors.

² For different reviews of the literature, see Basberg (1987), Griliches (1990) and Archibugi (1992).

³ European Commission, *Second European Report on Science & Technology Indicators*, 1997, EUR 17639, p.93.

present an intrinsic international nature while Triadic patent families are made of high value patents (e.g. Guellec and van Pottelsberghe (2004) and Sapsalis and van Pottelsberghe (2005)). In addition, the Triadic Patent Families suffer less from home advantage effects than USPTO or EPO patent data (Criscuolo (2005)).

The second methodological choice concerns the aggregation level. Technological classifications allow to link patents to one or more economic areas. There is no natural or perfect correspondence between technological classifications and economic areas. This problem of classification refers primarily to the difficulties in allocating patent data, organised by technological classes, into economically relevant industries or product groupings.⁴

European patents are classified according to the International Patent Classification (IPC) published by the World Intellectual Property Organisation (WIPO), a hierarchical classification with several levels of breakdown primarily concerned with the technological characteristics of the invention. Another classification, known as the OST-INPI/FhG-ISI technology nomenclature (or “OST” in short) and having two levels of industrial breakdown, is provided by the “Office des Sciences et Techniques”, the French Patent Office (INPI) and the Fraunhofer ISI Institute.

Each classification is organised in a hierarchical way, from generic and very large categories, to more detailed subcategories, and so forth. This means that besides the choice of a classification system, a choice has to be done regarding the level of aggregation. This is achieved by taking in account the appropriate number of digits in the patent’s classification code. For instance, if we consider only the first digit of the IPC classification system, we are at the uppermost level.⁵

In this paper, five different levels of classifications are considered:

- Level 1 of IPC (IPC1): 8 classes
- Level 3 of IPC (IPC3): 117 classes
- Level 4 of IPC (IPC4): 850+ classes
- Level 1 of OST (OST1): 6 classes
- Level 2 of OST (OST2): 30 classes

The third methodological choice concerns the statistical formula used to measure the technological specialisation of each country. In what follows, we focus on absolute concentration ratios, measuring the degree of specialisation of each country compared to the degree of specialisation of the others. Five different statistical measures are available:

- The concentration ratios with the 4, 8 or 20 biggest classes (C4, C8, C20) compute easily the share of the X biggest classes while neglecting the distribution between the X largest classes and the remaining classes.

⁴ See Griliches (1990) on the classification problem and Schmoch, Laville, Patel and Frietsch (2003).

⁵ As IPC codes are structured as “XNNY...” (for example, A16K...), one can consider only the first character of the code (A here), and classify patents according to the 8 generic classes, referring to the “IPC1” classification. If we take into account the two digits following the letter (A16), we can get up to 117 subclasses associated with “IPC3”. Finally, one can add up the fourth character and take into account a narrower level of classification, the “IPC4”, with more than 850 classes.

The C(X) ratio is given by $C(X)_j = \frac{\sum_{k=1}^X p_{(k)j}}{\sum_{i=1}^n p_{ij}}$ where X is the number of the largest classes to

take into account, p_{ij} is the number of patents (or applications) of country j in the i^{th} technological class with $i = 1, \dots, n$ where n is the total number of classes, and $p_{(k)j}$ is the k^{th} largest number of patents per technological class. The concentration ratios with the X largest classes requires therefore having more than X classes in total in the classification, which excludes IPC1 and OST1 with C8 or C20.

- The concentration index of Gini measures the deviation compared to a uniform distribution and is more sensitive to the presence of a large number of small classes.

The Gini coefficient is given by $G_j = 1 - \sum_{i=0}^{n-1} (F_{i+1,j} - F_{i,j})(\Phi_{i+1,j} + \Phi_{i,j})$

where $F_{i,j}$ is the cumulative population share and $\Phi_i = \frac{1}{n\mu} \sum_{k=1}^i p_{kj}$, the cumulative patent share of class i , p_{ij} being the number of patents (or applications) of country j in the i^{th} technological class with $i = 1, \dots, n$ where n is the total number of classes.

- The concentration measure of Herfindahl takes into account the whole distribution but the share of each class is weighted and the measure is less influenced by the presence of a big number of small classes. If s_{ij} is the share of the i^{th} technological class (with $i = 1, \dots, n$ where $n =$ total number of classes) in the total number of patents (or applications) of country j , the Herfindahl index is given by $H_j = \sum_{i=1}^n s_{ij}^2$.

3. The methodology

The degree of technological specialisation has been computed for each country, based on each data source, at every level of classification and with each statistical index for the two periods considered (1987-1989 and 1993-1995). This provides us the basis for the “static” analysis. The growth rate of each concentration measure between the two periods generates the “dynamic” analysis. This represents a total of 75 measures for each analysis (“static” and “dynamic”).

Based on these specialisation levels and growth rates, country rankings are established for each pair of aggregation level and statistical index, and for each data set, both in static and dynamic dimensions. In static terms, the most specialised countries according to a particular aggregation level, given a statistical index, and computed with one source of data, get the lowest (or best) rank. From a dynamic point of view, the countries with the highest growth rate in terms of specialisation between the two periods get the lowest rank. This method provides two specific country rankings (static and dynamic) for each of the 75 technological specialisation measures.

Table 1 provides several measures of technological specialisation using the C4 concentration index and USPTO patents for different aggregation levels during the period 1993-1995 and in terms of growth rates between 1987-89 and 1993-95.

***** Insert Table 1 about here *****

A sensitivity analysis can then be performed in three steps, to assess the influence of each of the three dimensions:

- In order to evaluate the impact of the aggregation level chosen, we compare the rankings obtained from each classification with different concentration indexes, with each data source independently, which can then also be pair wised compared.
- In order to evaluate the impact of the concentration index used, we compare the rankings obtained with each index using different aggregation levels and data sources.
- For the evaluation of the impact of the source of patent data, rankings based on the 3 different sources are compared for every possible aggregation level – concentration index combination (for example, Gini and IPC1, Gini and IPC3, etc.)

For two country rankings to be considered equivalent, the similarity measure must satisfy two conditions:

- The Spearman rank correlation ratio between two rankings needs to be significant at the one percent probability threshold. The Spearman rank correlation coefficient is defined by $r_s = 1 - \frac{6}{N(N^2 - 1)} \sum_{j=1}^N d_j^2$ where N is the number of countries in the ranking and d_j is the difference in rank of country j between the two rankings.
- The biggest rank difference among countries (or jump of each country) between two rankings must be no bigger than a particular threshold (T), or $d_j < T \quad \forall j = 1, \dots, N$. Since a threshold of 4 would not have the same effect on rankings of 15 or 10 countries, it has to be determined in relative terms, as a particular percentage of N , the number of countries in the rankings. A threshold of 25% for example would become 4 with 16 countries in the rankings and 3 with 11 countries in the rankings.⁶

***** Insert Table 2 about here *****

An example of pair wise similarity test results is presented in Table 2. Ultimately, for each possible choice along every dimension, we add-up the number of equivalent ranking pairs, computed with every possible combination of choices along the two other dimensions. For instance, in Table 2, only one pair of rankings (in static terms, IPC3/OST1) would be considered similar out of 10 with a threshold of 33%. A high number of similar ranking pairs would denote a low sensitivity of the measure to the choice of an aggregation level. For instance, if most of the rankings computed from EPO data with the

⁶ In relative terms, our results are not strongly affected by a change in the relative thresholds, but it appears that the number of pairs of similar specialisation rankings following the similarity test is the largest with a threshold of 50% and is not significantly improved with higher values of the threshold. This means that most pairs of rankings satisfying the Spearman correlation test condition present maximum rank differences that do not exceed 50% of the number of countries. However, 50% represents an important jump in the rankings, and 25% or 33% seem more appropriate and relevant values to consider the rankings as similar.

IPC3 classification but using 5 different concentration indexes were similar, it would mean that the concentration index has little impact on the measure and is therefore not determinant or critical in the measurement of technological specialisation.

4. The data and statistical results

For each of the data sources (patents granted by the USPTO, patent applications at the EPO, Triadic patent families), we consider two periods, 1987-1989 and 1993-1995, for different countries (16 countries with USPTO and EPO data and 11 with Triadic patent families). Data from the period 1993-1995 is used for the static analysis, while the growth rates in specialisation between the two periods is used for the dynamic analysis.

In each data set, we count the number of patents of each country by technological class for the covered period, according to the country of residence of the inventor(s) and to the main IPC class of the patent or application.

The main results of this study can be presented in terms of the impact of each dimension of the measure on the country rankings. This impact is assessed in a summary table showing the number of pairs of similar specialisation rankings for each set of measurement options at a threshold of 25%.

***** Insert Table 3 about here *****

Regarding the impact of the technological classification, Table 3 shows an extremely low level of similarity between the pairs of rankings with different aggregation levels. With a maximum rank difference of 25%, there are virtually no equivalent pairs of rankings. The impact is stronger with USPTO and Triadic data. These results show that the choice of a technological aggregation level has a strong impact on country rankings. Rankings computed with the C20 or Gini based on EPO data seem slightly less affected.

The impact of the concentration index is presented in Table 4. This choice is less significant than the choice of an aggregation level. Even at a threshold of 25%, more than half of the pairs compared pass our similarity test on static rankings, and one third on dynamic comparisons. For instance, taking the index computed with USPTO data and an aggregation level of IPC-3, table 4 shows that 90% of the country ranking pairs across the various statistical measures are similar. The similarities are even close to perfect if we tolerate a maximum rank difference of 50%, meaning that almost every pair passes the Spearman correlation test and show a maximum jump of less than 50% of the number of countries in the ranking.

***** Insert Table 4 about here *****

These results witness a very weak impact of the concentration indexes on the specialisation rankings, although slightly stronger on dynamic rankings than on static rankings. Additionally, more stable and coherent results in both static and dynamic specialisation rankings seem to be obtained with highest aggregation levels (levels with more classes, i.e. IPC3 and IPC4) and with EPO data.

Table 5 shows that the choice of the data source seems to have a serious impact on technological specialisation rankings. It is however interesting to notice that the impact is much stronger on static specialisation rankings than on dynamic specialisation rankings. Indeed, there is virtually no pair of

rankings passing the similarity test with a threshold of 25%. These observations are not altered by the use of higher thresholds.

***** Insert Table 5 about here *****

These results suggest that the choice of a data source has a very strong impact on specialisation rankings, although this impact can be significantly reduced on dynamic rankings when using specific concentration indexes and specific classifications. This is particularly the case of the IPC3 and IPC4 classifications and of the C8, C20 and Gini concentration ratios, which confirms our previous observations.

Table 6 summarises the overall sensitivity of each option in one dimension to the other dimensions. This table shows first that technological specialisation measures are much more sensitive to data sources and aggregation levels than to statistical indexes. It confirms also that contrarily to data sources, concentration indexes and aggregation levels have more impact on dynamic than on static measures. One can finally observe that measures based on IPC3 and IPC4 are relatively less sensitive to choices along the other dimensions, as are measures using the coefficient of Gini or a concentration index with the 20 biggest classes and computations with EPO applications data.

***** Insert Table 6 about here *****

5. Concluding remarks

When measuring technological specialisation with patent data, researchers must be extremely careful when choosing their options along the three dimensions of the measure. Though the effect of the choice of a concentration measure is moderated, the Gini and C20 indexes are probably the most reliable choices. Of utmost importance, the aggregation level of the classification must be chosen so as to have a large number of classes, meaning at least the third level of IPC. However, given the difficulty to match IPC3 classes to economic sectors, IPC4 is certainly more appropriate.⁷

Finally, if the choice of a data source critically determines the degree of specialisation of countries, it might not always capture the same specialisation due mainly to the home advantage bias.⁸ When measuring the degree of specialisation of a country using USPTO grants, researchers should be aware that such a measure would provide the overall degree of specialisation of the United States, but probably a more restrictive degree of specialisation of European or Japanese patents, since only European and Japanese patents of higher value are also filed at USPTO. This could explain why measures based on EPO data are on average more stable than those using USPTO data.

Triadic Patent Families could therefore provide the most neutral data source, but since they only take into account patents of sufficient value as to be filed simultaneously in EPO, JPO and USPTO, measures based on such patent families would again provide a truncated view of a country's degree of technological specialisation and appear empirically less stable.

⁷ See Schmoch, Laville, Patel & Frietsch (2003).

⁸ See van Leuven (1996) and Criscuolo (2005).

References

- ARCHIBUGI D. (1992), “*Patenting as an Indicator of Technological Innovation: A Review*”, Science and Public Policy, Vol.19, No.6, pp. 357-358.
- ARCHIBUGI D. and PIANTA M. (1992), “*The Technological Specialization of Advanced Countries*”, Report to the EEC on International Science and Technology Activities (Kluwer, Boston).
- BASBERG B.L. (1987), “*Patents and the Measurement of Technological Change: A Survey of the Literature*”, Research Policy, 16, pp. 131-141.
- BRAUN T., GLANZEL W. and GRUPP H. (1995), “*The Scientometric Weight of 50 Nations in 27 Science Areas*”, 1989-1993. “Part I. All Fields Combined, Mathematics, Engineering, Chemistry and Physics”, Scientometrics, 33, pp. 263-293; “Part II. Life Sciences”, Scientometrics, 34, pp. 207-237.
- COMANOR W.S. and SCHERER F.M. (1969), “*Patents Statistics as a Measure of Technical Change*”, Journal of Political Economy, Vol.77, pp.392-398.
- CRISCUOLO P. (2005), “*The ‘home advantage’ effect and patent families: a comparison of OECD triadic patents, the USPTO and the EPO*”, miméo, Tanaka Business School, Imperial College London. Paper presented at the EPIP Conference, Copenhagen, 10th and 11th March 2005.
- DERNIS H., GUELLEC D. and VAN POTTELSBERGHE B. (2001), “*Using patent counts for cross-country comparisons of technology output*”, STI Review, 27, OECD, Paris, pp. 129-146.
- DERNIS H. and KHAN M. (2004), “*Triadic Patent Families Methodology*”, STI Working Paper 2004/2, OECD, Paris.
- EPO, JPO and USPTO, “*Trilateral Statistical Report, 1995*”.
- EUROPEAN COMMISSION (1997), “*European Report on S&T Indicators 1997*”, EUR 17639, pp.93 and 99-106; and (2003), “*European Report on S&T Indicators 2003*”, EUR 20025, pp.294-300.
- GRILICHES Z. (1990), “*Patent statistics as economic indicators: A Survey*”, Journal of Economic Literature, 28(4), 1661-1707.
- GUELLEC D. and B. VAN POTTELSBERGHE (2004), “*Measuring the internationalisation of the generation of knowledge - An approach based on patent data*” in Chapter 29, jointly with D. Guellec, in H.F. Moed, W. Glänzel and U. Schmoch (Eds), Handbook of Quantitative Science and Technology Research – The use of publication and patent statistics in studies of science and technology systems, Kluwer Academic Publishers, 2004, pp. 645-662.
- OBSERVATOIRE DES SCIENCES ET TECHNIQUES (1993), “*Science et Technologie Indicateurs 1994*”, Economica, Paris.
- OECD (1994), “*The Measurement of Scientific and Technological Activities. Using Patent Data as Science and Technology Indicators – Patent Manual*”, OECD, Paris, 1994.

SAPSALIS E. and B. VAN POTTELSBERGHE (2005), “*The sources of knowledge and the value of academic patents*”, Economics of Innovation and New Technologies, forthcoming.

SCHMOCH U., LAVILLE F., PATEL P. and FRIETSCH R. (2003), “*Linking technology areas to industrial sectors*”, Final Report to the EC Commission, DG Research, Karlsruhe – Paris – Brighton.

SCHMOOKLER J. (1972), “*Patents, Invention and Economic Change*”, Harvard University Press, Cambridge, Massachussets.

VAN LEUVEN J. (1996), “*Patent Statistics as indicators for innovation*”, Patent World, November/December 1996, pages 20-27.

WORLD INTELLECTUAL PROPERTY ORGANISATION (2000), “*International Patent Classification (IPC)*”, <http://www.wipo.int/classifications/en/>

Tables

Table 1: Degree of Specialization by country (and rank of the country) calculated with C4 and USPTO data

	1993-1995					Growth rate (1993-1995 v/s 1987-1989)				
	IPC1	IPC3	IPC4	OST1	OST2	IPC1	IPC3	IPC4	OST1	OST2
Germany	.71 (15)	.26 (13)	.11 (16)	.82 (08)	.27 (16)	-2% (15)	3% (12)	-3% (14)	0% (10)	2% (09)
Australia	.74 (10)	.25 (14)	.13 (12)	.74 (15)	.29 (14)	3% (05)	-6% (16)	1% (11)	-2% (14)	-2% (12)
Austria	.71 (13)	.22 (16)	.13 (13)	.78 (12)	.31 (09)	-1% (14)	-1% (14)	-2% (12)	1% (06)	4% (06)
Belgium	.81 (04)	.35 (06)	.22 (03)	.84 (03)	.37 (04)	3% (04)	3% (10)	27% (04)	0% (12)	-2% (11)
Canada	.71 (16)	.24 (15)	.12 (15)	.71 (16)	.28 (15)	-1% (13)	18% (04)	31% (03)	-4% (16)	-13% (16)
South Korea	.86 (01)	.49 (01)	.33 (01)	.87 (01)	.49 (01)	1% (08)	3% (11)	6% (09)	1% (05)	-4% (13)
Denmark	.83 (03)	.41 (02)	.24 (02)	.84 (04)	.36 (05)	5% (03)	7% (08)	1% (10)	1% (09)	7% (05)
United States	.76 (06)	.30 (12)	.16 (08)	.76 (14)	.29 (11)	6% (02)	19% (03)	46% (01)	4% (01)	3% (07)
Finland	.72 (11)	.38 (03)	.20 (04)	.83 (05)	.45 (02)	10% (01)	33% (01)	37% (02)	2% (04)	13% (02)
France	.71 (14)	.32 (09)	.17 (07)	.80 (10)	.29 (11)	2% (07)	9% (07)	24% (06)	1% (07)	9% (04)
Italy	.76 (07)	.31 (10)	.15 (10)	.80 (10)	.31 (08)	1% (09)	-3% (15)	-6% (16)	0% (11)	-6% (14)
Japan	.85 (02)	.35 (05)	.20 (05)	.87 (02)	.39 (03)	2% (06)	17% (05)	12% (08)	2% (02)	12% (03)
Netherlands	.76 (05)	.32 (08)	.13 (14)	.82 (07)	.31 (10)	-3% (16)	12% (06)	-5% (15)	-2% (15)	-1% (10)
United Kingdom	.72 (12)	.31 (11)	.18 (06)	.77 (13)	.29 (13)	-1% (12)	4% (09)	15% (07)	1% (08)	3% (08)
Sweden	.75 (09)	.36 (04)	.14 (11)	.81 (09)	.34 (06)	0% (11)	25% (02)	25% (05)	2% (03)	18% (01)
Switzerland	.75 (08)	.32 (07)	.15 (09)	.82 (06)	.31 (07)	0% (10)	2% (13)	-2% (13)	-2% (13)	-11% (15)

Table 2: Maximum rank differences and Spearman independence for country rankings with different classifications calculated with C4 and with USPTO data

	1993-1995					Growth rate (1993-1995 v/s 1987-1989)				
	IPC1	IPC3	IPC4	OST1	OST2	IPC1	IPC3	IPC4	OST1	OST2
IPC-1	0	-	-	-	-	0	-	-	-	-
IPC-3	8 (*)	0	-	-	-	11	0	-	-	-
IPC-4	9	7 (*)	0	-	-	10	9 (*)	0	-	-
OST-1	8 (*)	5 (*)	8 (*)	0	-	9	12	13	0	-
OST-2	9 (*)	7 (*)	7 (*)	8 (*)	0	10	12	13	8 (*)	0

(*) = Pair of technological specialization rankings showing a significant correlation (Spearman) at the level of 1%

Table 3: Number of pairs (and percentages) of similar specialization rankings according to different classifications. Pairs added from measures of each data source with a threshold (T) of 25%

	USPTO		EPO		TRIAD		Total		Number of pairs
	93-95	Growth	93-95	Growth	93-95	Growth	93-95	Growth	
C4	0 (00%)	0 (00%)	0 (00%)	0 (00%)	1 (10%)	0 (00%)	1 (03%)	0 (00%)	10
C8	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	3
C20	0 (00%)	0 (00%)	1 (33%)	1 (33%)	0 (00%)	0 (00%)	1 (11%)	1 (11%)	3
Herfindahl	0 (00%)	0 (00%)	1 (10%)	2 (20%)	0 (00%)	2 (20%)	1 (03%)	4 (13%)	10
Gini	1 (10%)	1 (10%)	1 (10%)	1 (10%)	1 (10%)	0 (00%)	3 (10%)	2 (07%)	10
Total	1 (03%)	1 (03%)	3 (08%)	4 (11%)	2 (06%)	2 (06%)	6 (06%)	7 (06%)	36

Table 4: Number of pairs (and percentages) of similar specialization rankings according to different ratios. Pairs added from measures of each data source with a threshold (T) of 25%

	USPTO		EPO		TRIAD		Total		Number of pairs
	93-95	Growth	93-95	Growth	93-95	Growth	93-95	Growth	
IPC1	2 (67%)	1 (33%)	2 (67%)	1 (33%)	3 (100%)	0 (00%)	7 (78%)	2 (22%)	3
IPC3	9 (90%)	2 (20%)	4 (40%)	6 (60%)	1 (10%)	3 (30%)	14 (47%)	11 (37%)	10
IPC4	5 (50%)	1 (10%)	10 (100%)	6 (60%)	4 (40%)	6 (60%)	19 (63%)	13 (43%)	10
OST1	3 (100%)	1 (33%)	1 (33%)	1 (33%)	1 (33%)	0 (00%)	5 (56%)	2 (22%)	3
OST2	6 (60%)	0 (00%)	5 (50%)	3 (30%)	6 (60%)	1 (10%)	17 (57%)	4 (13%)	10
Total	25 (69%)	5 (14%)	22 (61%)	17 (47%)	15 (42%)	10 (28%)	62 (57%)	32 (30%)	36

Table 5: Number of pairs (and percentages) of similar specialization rankings of 25% according to different data sources with a threshold (T)

	IPC1		IPC3		IPC4		OST1		OST2		Total		Nb of pairs
	93-95	Growth	93-95	Growth	93-95	Growth	93-95	Growth	93-95	Growth	93-95	Growth	
C4	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	3
C8	-	-	0 (00%)	1 (33%)	0 (00%)	1 (33%)	-	-	0 (00%)	0 (00%)	0 (00%)	2 (13%)	3
C20	-	-	0 (00%)	2 (67%)	0 (00%)	1 (33%)	-	-	0 (00%)	0 (00%)	0 (00%)	3 (20%)	3
Herfindahl	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	1 (33%)	0 (00%)	1 (07%)	3
Gini	0 (00%)	0 (00%)	0 (00%)	1 (33%)	0 (00%)	3 (100%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	0 (00%)	4 (27%)	3
Total	0 (00%)	0 (00%)	0 (00%)	4 (27%)	0 (00%)	5 (33%)	0 (00%)	0 (00%)	0 (00%)	1 (07%)	0 (00%)	10 (13%)	15

Table 6: Overall sensitivity of each choice along each dimension to the other dimensions

	Classification level		Concentration index		Data source		Overall Sensitivity
	Static	Dynamic	Static	Dynamic	Static	Dynamic	
IPC1	-	-	LOW	MEDIUM	HIGHEST	HIGHEST	HIGH
IPC3	-	-	LOW	LOW	HIGHEST	MEDIUM	MEDIUM
IPC4	-	-	LOW	LOW	HIGHEST	MEDIUM	MEDIUM
OST1	-	-	LOW	MEDIUM	HIGHEST	HIGHEST	HIGH
OST2	-	-	LOW	MEDIUM	HIGHEST	HIGHEST	HIGH
C4	HIGHEST	HIGHEST	-	-	HIGHEST	HIGHEST	HIGHEST
C8	HIGHEST	HIGH	-	-	HIGHEST	HIGH	HIGHEST
C20	HIGH	HIGH	-	-	HIGHEST	HIGH	HIGH
Herfindahl	HIGH	HIGHEST	-	-	HIGHEST	HIGH	HIGHEST
Gini	HIGH	HIGHEST	-	-	HIGHEST	MEDIUM	HIGH
USPTO	MEDIUM	HIGHEST	LOWEST	MEDIUM	-	-	MEDIUM
EPO	HIGH	HIGH	LOWEST	LOW	-	-	MEDIUM
Triadic	HIGHEST	HIGHEST	MEDIUM	MEDIUM	-	-	HIGH
Overall Impact	HIGH	HIGHEST	LOW	MEDIUM	HIGHEST	HIGH	-
	HIGH		LOW		HIGHEST		-