Linking Technology Areas to Industrial Sectors[1]

Ulrich Schmoch[2]
Francoise Laville[3]
Pari Patel[4]
Rainer Frietsch[1]

1. Introduction

Technological innovations are one of the key factors in explaining economic competitiveness of advanced countries. Therefore it is important to monitor technological development of areas, countries and regions in a systematic way to support economic analysis and decision making. It is, however, impossible to describe the technological development by a single indicator encompassing all aspects and stages of innovation. Rather, it is necessary to establish a network of related indicators reflecting different aspects of innovation. At the same time to examine the relationship between technology and economic performance it is crucial to link technological indicators with those related to economic performance. At the international level, most economic indicators such as turnover, investment, employment, productivity, value added, R&D expenditure etc. are classified by industrial sectors, for instance, according to the NACE or ISIC schemes. In contrast, some of the most frequently used indicators for technology are based on patent statistics, classified according to the International Patent Classification (IPC).[5] However, the IPC is based on technological categories and cannot be directly translated into industrial sectors. One approach for solving this problem is to establish a reliable concordance between technology and industry classifications.

2. State of the art

There have been a number of attempts in the past to establish a link between technological and economic indicators. However these concordances have not found satisfactory solutions to the following four problems: (1) international comparability, (2) level of disaggregation (3) strong empirical basis, (4) easy applicability to specific problems.

1 The research underlying this paper was supported by the European Commission, DG Research.

2 Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany, e-mail: us@isi.fhg.de

3 Observatoire des Sciences et des Techniques (OST), Paris, France

4 University of Sussex, Science and Policy Research Unit (SPRU), Brighton, United Kindom

5 The relevance of national patent classifications such as the USPOC is decreasing, as in all cases, the national patents are additionally classified in terms of the IPC.

Furthermore, since some of these were established, industrial structures have changed, necessitating a change in the nomenclatures.

One of the earliest attempts at linking technology and industry classifications was by Kronz (1980), who classified the patent applications of four countries by NACE classes. This was based more on an intuitive approach, rather than on the basis of a systematic analysis leading to a well-defined concordance table.

Evenson / Puttnam (1988) use data from the Canadian Patent Office, where patent examiners simultaneously assigned IPC codes together with an industry of manufacture and sector of use to each of 300,000 patents granted between 1972 and 1995. On the basis of these data, they established a cross-tabulation between 8 IPC sections and 25 industries, called the Yale-Canada patent flow concordance. The two main problems with this approach, which limit its value in terms of practical applications are: (a) it is based on Canadian SIC, which needs to be translated to either ISIC Rev 3 or NACE; and (b) it is not very detailed in terms of IPC codes. An additional difficulty is that the relationship between sectors and technologies has distinctly changed during the period 1972 to 1995.

Verspagen et al. (1994) suggested a concordance scheme between four-digit level IPC subclasses and 22 (2 and 3 digit) industrial classes based on ISIC (rev. 2), the so-called MERIT Concordance. The linkage was established by an intellectual approach, and based on a similar concordance of Statistics Finland. In this approach, many of the 625 IPC subclasses are linked with different weights to different sectors, so that it is quite time-consuming to calculate statistics for specific sectors.

In the 1980s the US Patent and Trademark Office established a detailed concordance between subclasses of the USPC and 41 unique classes of the US Standard Industrial classification, and this is used to produce regular statistics of US patents by SIC sectors. This is simply done on the basis of examining the definition of each USPC class (and sometimes subclass) and assigning them to one or more of the 41 industrial classes. For our purposes this concordance has some of the problems already identified above. It is based on the USPC and not the IPC, limiting its applicability to EPO data. Further the industrial classification used is the US-SIC, which needs to be translated into ISIC for practical use.

Greif / Potkowik (1990) computed statistics of patents by industrial sectors based on an old German national statistical classification scheme (Wirtschaftszweige, WZ79) which is not compatible with the present NACE or ISIC codes. They assigned WZ codes to a sample of 280 applicants in 1983 at the German Patent Office and analysed their patent activities in terms of IPC codes. Again the validity for present purposes is quite limited.

The most recent attempt at defining a concordance between IPC and ISIC codes is by Johnson (2002). As with the earlier work of Evenson / Putnam (1988, see above), this is based on data from the Canadian Patent Office. For 625 IPC subclasses, Johnson defines probabilities of linkages to different sectors of manufacture and use. However, this interesting method has several limitations. Firstly the linkage between IPC codes and sectors is defined by examiners of the Canadian Intellectual Property Office, and is not based on the official industrial class of the company to whom the patent is assigned. This is likely to result in a strong technology bias. Secondly the Canadian Office stopped assigning sector codes to patents in the grant year 1995, equivalent to about 1991 in terms of first application (priority). Thus, the concordance is quite old, and there is a high probability that the relationship between technology and sectors has changed since then. Thirdly the sectors are defined in terms of Canadian SIC codes, and have to be translated into ISIC codes, implying certain inaccuracies due to translation. Fourthly the concordance is based on the determination of 70,000 probabilities of linkage between IPC and ISIC codes. Therefore it can only be handled with the support of a complex software package, consisting of three separate modules. Moreover as input, the user has to provide search results for all IPC subclasses which requires access to a comprehensive large-scale patent database.

To sum up, there is still a need to design a concordance which provides a linkage between sector and technology classifications, and which can be handled in a straightforward way.

3. Development of a basic association of technologies and sectors

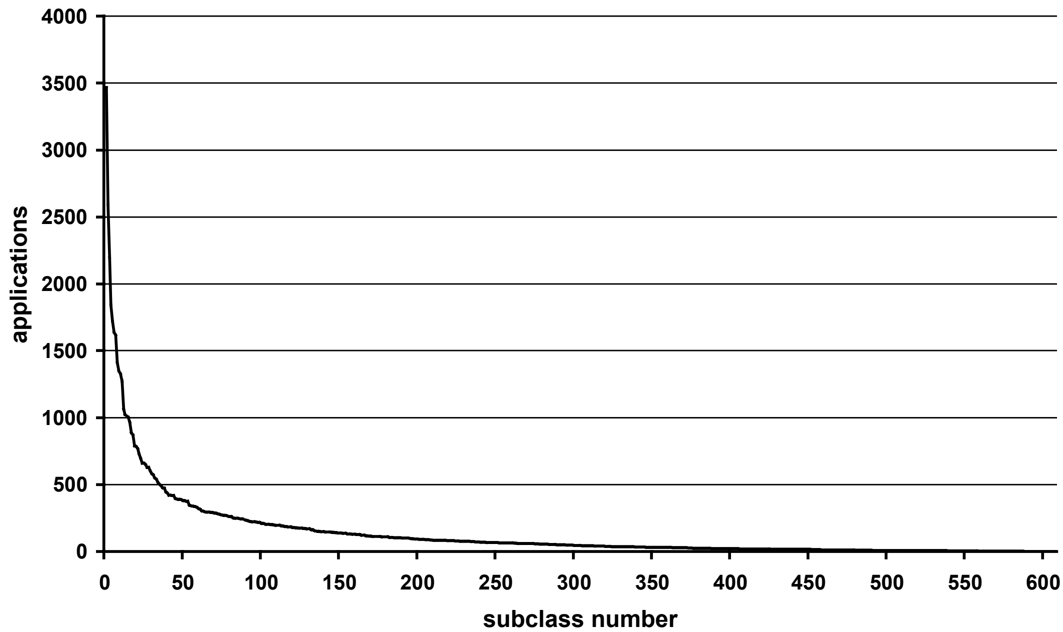
Our approach starts with the selection of industrial sectors at the 2-digit level of NACE or ISIC with a finer breakdown of the quantitatively important sectors within chemicals, machinery, and electrical equipment, leading to 44 sectors of manufacture, documented in annex 1. This level of disaggregation is finer than most statistics on economic data, e.g. foreign trade, value added, or R&D expenditure, provided by OECD, Eurostat or other authorities. It was chosen to be able to show the main differences between the sub-sectors in chemicals, machinery, and electrical equipment industries. Thus a higher level of aggregation can be achieved by a simple combination (addition) of sub-sectors. Moreover it is possible to transfer the NACE-defined fields directly into ISIC-based sectors.

Industrial sectors are defined by the manufacturing characteristic of products, so that it is possible to associate them to technologies. On this basis, technical experts of Fraunhofer ISI associated each of the 625 subclasses of the IPC to one of the industrial categories mentioned above. In the following text, this first step is called "original, intellectual association" of technologies and sectors. Only subclasses that appear as primary classes were used, ignoring all index/cross reference codes. In the case of the subclasses F21M to F21Q (lighting), we included older codes that do not appear in the latest version (7th version) of the IPC. By this means, it will be possible to compute longer time series starting in about 1985.

The IPC subclasses were linked to one field only, even if multiple linkages to other fields were obvious, by applying the principle of main focus. In unclear cases, we made a statistical check of secondary IPC codes which generally led to a clear decision. In the few cases[6] where picture was unclear, we decided not to split the classes into different sectors, in order to keep the basic structures sufficiently clear.

The impact of the most "unclear" cases is generally low in terms of absolute numbers. The volume of patenting by IPC subclasses is quite uneven, as shown in figure 1. In the largest subclass, 3469 patent applications were registered in 1997, whereas in many subclasses no applications were filed at all. The top ten percent of subclasses (a segment of about 60 subclasses), account for 58 percent of all applications, and the top 30 percent for 85 percent (cf. figure 2) of the total. This implies that only the largest unambiguous cases would need to be considered in more detail. As an example, the subclass H03K (Pulse technique) is linked nearly in equal parts to telecommunications and computer technology and is sufficiently large to deserve special attention, with about 300 EPO applications in one year.

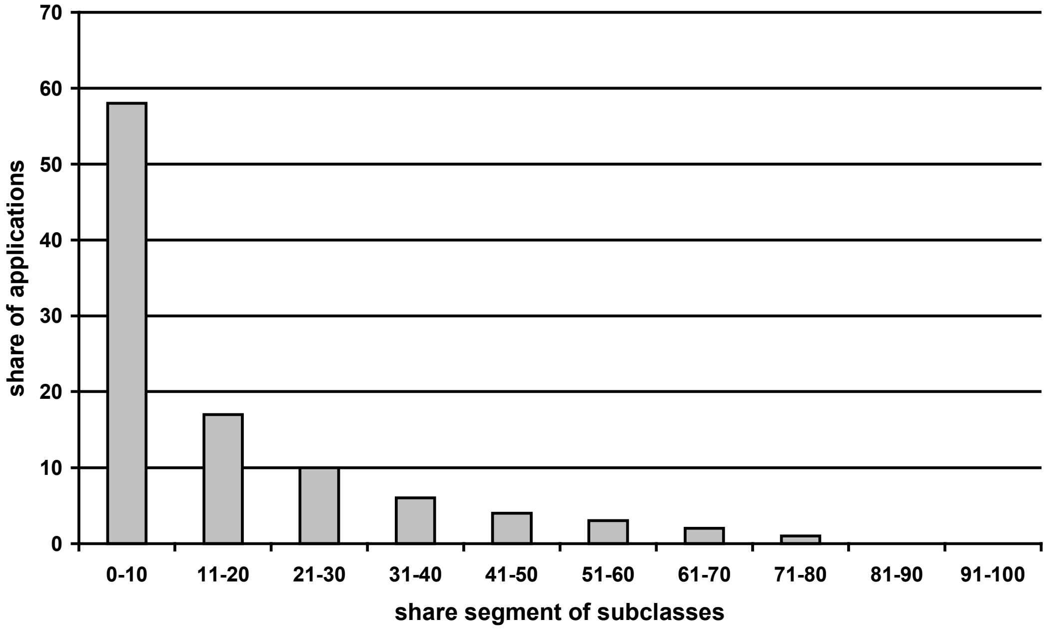
Figure 1: Distribution of the number of EPO applications by IPC subclasses in the priority year 1997



Source: EPAT; Ccalculations of Fraunhofer ISI

6 E. g. H03K, Pulse technique, or some subclasses of C07, Organic chemistry.

Figure 2: Distribution of the number of EPO applications by segments of IPC subclasses (shares in percent, ordered by application volume) in the priority year 1997



Source: EPAT; Computation by Fraunhofer ISI

4. Empirical basis of verification

The original association of technologies and sectors was exclusively defined on the basis of expert assessment, and needed a further empirical verification. This was achieved using an offline database of OST[7] containing all the data on European and PCT applications without double counting. The information for each patent includes IPC codes, inventors, and applicants with geographical information. This was supplemented by data from Dun & Bradstreet (D&B) which assisted in classifying each applicant by industry. In the D&B database, the industrial activities of firms are described using the US SIC classification, so that they had to be transferred to NACE codes for the purpose of the current project. Although there is no exact correspondence between SIC and NACE codes, at a low level of aggregation (such as the 44 classes mentioned above) it is possible to establish a good association between the classifications. To sum up, the match of SIC and NACE (ISIC) is reliable at this high level of aggregation .

7 Oberservatoire des Sciences et des Techniques, Paris.

OST only conducted this exercise for institutional applicants with more than 5 patent applications in three years, called "large firms". For the period 1997 to 1999, the OST database has more than 50,000 applicants. However only 3,900 applicants registered more than 5 patents in three years, and are thus considered as "large firms". These 3900 applicants account for more than 65 % of all patents applied for by institutions.

The first steps of the empirical verification were based on the large firms' data set for the publication period 1997 to 1999, exclusively using the first, main IPC codes (IPC_1). This data set comprises 2450 enterprises. For the period 1990 to 1992, the analysis is based on 2080 firms. Furthermore, the three partner institutes Fraunhofer ISI, OST, and SPRU generated a second data set of "small firms" (defined as those with less than 5 applications in 3 years), located in France, Germany and the United Kingdom. For these firms, the industrial sector codes were identified by national data sources and integrated into the OST database. The data set of small firms began with about 3,000 applicants, and it was possible to assign NACE codes in about 2,360 cases.

On this basis, the original intellectual association of patent and industrial classifications was improved by information on industrial sectors of both the small and large firms registering patents in each IPC subclass. This second step is called "revised association". In the case of small firms, the association of sector and IPC codes proved to be closely linked to the technological content, whereas the technological orientation is blurred for large firms due to their broader spectrum of industrial activities. However, the number of patents in many IPC subclasses were too small for the small firms, so that the results for larger firms were necessary to decide on an appropriate association. The final decision on the association of subclasses was taken intellectually, as the orientation of the concordance should be primarily technological. In some technologies patents are not primarily taken out by firms belonging to the industrial sector which could in some sense be regarded as being "responsible" for this technology. For instance, the enterprises in "Basic chemicals"[8] are often the most frequent applicants in IPC classes linked to "Other chemicals"[9] in terms of technology. In these cases, the decision was taken in favour of the technological content, if the "responsible" sector has still a substantial number of applications in the IPC subclass considered.

Table 1 illustrates the approach for this additional empirical check. For instance, most patent applications in the subclass A01B belonged to firms classified in industrial field 23 (see column 1 in the Table in Annex 1). This is in agreement with the original intellectual association. The same applies to the IPC code A01C. In the case of the code A01G, most patents were assigned to firms belonging to industrial field 23 with a nearly equivalent level for those in sector 9. In the original, intellectual association, this IPC code was linked to field 20 which appears to be less relevant, as indicated in Table 1.

8 NACE code 24.1.

9 NACE code 24.6.

The code A01G is "Horticulture, cultivation of vegetables, watering", industrial field 20 is "Fabricated metal products". Thus in the intellectual association, the focus of patent applications was wrongly assumed to be on tools for Horticulture. With the additional information the choice is to redefine the association of code A01G to sector 23 (Agricultural machinery), or sector 9 (Petroleum products). All in all, a revised linkage of A01G to field 23 seems to be appropriate, but this choice is of course ambiguous and a matter of judgement.

Table 1: Patent applications in IPC subclasses by industry-based fields for large firms

IPC	Field	#	Share
A01B	23	53	98.1
	42	1	1.9
A01C	23	62	84.9
	10	3	4.1
	16	3	4.1
	13	2	2.7
	11	1	1.4
	3	1	1.4
	7	1	1.4
A01G	23	10	19.2
	9	7	13.5
	1	6	11.5
	13	6	11.5
	42	5	9.6
	18	3	5.8
	19	3	5.8
	10	2	3.8
	11	2	3.8
	14	1	1.9
	15	1	1.9
	16	1	1.9
	20	1	1.9
	25	1	1.9
	27	1	1.9
	39	1	1.9
	7	1	1.9

Source: Computations of OST and Fraunhofer ISI

The linkage of the 625 IPC subclasses to the 44 fields would be different from the original intellectual association in about 60 percent of the cases, if the most important sector of patenting was taken as criterion. By considering in addition the patent activities of small firms and the technological content, the necessary changes are less frequent. However the original association was still amended in about 30 percent of the cases. So the information on the firms involved in each patent class resulted in a considerable refinement and improvement of the pure intellectual concordance.

5. Generation of a concordance matrix

The revised association has a technological orientation and should be labelled "ideal", as in reality, many firms provide a broader spectrum of products than indicated by the sector definitions. If the industrial sectors and the associated technology areas were in exact agreement, only the diagonal elements of a cross-tabulation in a matrix of 44 technological fields and 44 industrial fields would be filled. Table 2 illustrates the structure of such a matrix with seven fields. In the case of complete equivalence between technologies and industries, all applications should appear as diagonal elements D. However the results of the empirical analysis show that this is not the case, as there are a substantial number of patents in the non-diagonal fields. Several reasons may play a role:

- The linkage of an IPC code to a sector is "wrongly" assigned, i. e. the IPC code refers to a product range, not covered by the industrial sector
- The technology field cannot be linked to one sector in an unambiguous way, but it is linked to several sectors. The aim of the revised association, in any way, is to link the field to the most relevant sector for this technology.
- The firms in a sector are active in several technologies, partly because they are large multi-product firms, and partly because the products they produce are multi-technology. In the case of smaller firms, the concentration of applications on the diagonal of the matrix should be stronger.

Table 2: Linkage structure between technological fields and industrial sectors presented in matrix form

		Technological field						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Industrial fields	**1**	**D1**	N11	N12	N13	N14	N15	N16
	2	N21	**D2**	N22	N23	N24	N25	N26
	3	N31	N32	**D3**	N33	N34	N35	N36
	4	N41	N42	N43	**D4**	N44	N45	N46
	5	N51	N52	N53	N54	**D5**	N55	N56
	6	N61	N62	N63	N64	N65	**D6**	N66
	7	N71	N72	N73	N74	N75	N76	**D7**

The linkage of an IPC subclass to a "wrong" field was reduced to a minimum by the empirical checks described in section 4. The multiple association of an IPC code to dif-

ferent fields due to the multiple use of technologies or the multiple activities of companies show the technological interconnection of different fields, thus industrial sectors. This is the conceptual reason why the association of IPC codes to sectors was not purely made by a statistical choice of the most important industry in terms of patent applications. Furthermore the IPC codes were exclusively associated to one field, so that interconnections between different sectors become visible in the concordance matrix.
For generating the concordance matrix, the following decisions were made:

- Only large firms are considered. Due to the strong concentration of applications on these large actors (see section 4), this restriction seems to be justified. In country comparisons, the impact of small firms is less relevant.
- Only firms of the manufacturing sector are taken into account, as share of the service sector in all patent activities is about 3 percent (Blind et al. 2003).
- Only the principle industrial activity of a firm is taken into account, although many large firms have activities spread across different sectors, sometimes even in the service sectors. The main reason for this choice is the lack of precise information on the distribution of production or sales of firms across different industries. Besides, other statistics on economic data are based on the same principle.
- Only the first IPC code is taken into account. Previous comparisons have shown that the distributions of patents according to the first classification only and those according to first and secondary classifications are quite similar (Schmoch et al. 1988). However, more distinct differences may appear in the case of pharmacy, as many patents relevant for this area have first codes in the chemical area and only a secondary code in Pharmaceuticals (A61K, A61P). This effect will be less important for the present context, as the major relevant chemical subclasses are already included in the definition of field 13 (Pharmaceuticals).

Table 3 shows the outcome of the analysis for the first seven fields, i.e., this is an extract of the full 44x44 matrix. In most cases, the diagonal elements receive the highest number of applications. However this is not the case in technical field 3 (Textiles), where the industrial field / sector 7 (Paper) has a large number of applications.

The focus on diagonal elements, i. e., the level of correspondence between technical and industrial field definitions, can be visualised in two ways. Table 4 shows the distribution of applications within a technological field across industrial sectors. In field 1, the correspondence between technology and sector is high, whereas in the technical field 4 (Wearing), a considerable share of applications is taken out by the industrial fields / sectors 3 and 7 (Textiles and Paper). However, the comparison with table 3, according to the technological activities of the sectors, shows that this is largely due to the effect of low absolute application numbers of the Wearing sector. Any "irregularities" in the patent activities of some firms become more visible in small sectors, as they are not counterbalanced by the patents of the "normal" firms.

Table 3: Absolute application numbers for large firms (extraction for the first seven NACE sectors)

industrial sectors		technological fields: food 01	tobacco 02	textiles 03	wearing 04	leather 05	wood products 06	paper 07
food	**01**	690	0	10	0	0	0	4
tobacco	**02**	0	23	0	0	0	0	0
textiles	**03**	0	4	57	7	0	2	7
wearing	**04**	0	0	4	6	5	0	0
leather	**05**	0	0	8	1	29	0	0
wood products	**06**	0	0	1	0	0	18	3
paper	**07**	1	1	28	6	0	0	199

Source: Computation of OST

Table 4: Share of industrial sectors within technological fields with reference to the diagonal elements (vertical comparison in %) (extraction for the first seven NACE sectors)

industrial sectors		technological fields: food 01	tobacco 02	textiles 03	wearing 04	leather 05	wood prod-ucts 06	paper 07
food	**01**	100	0	18	0	0	0	2
tobacco	**02**	0	100	0	0	0	0	0
textiles	**03**	0	17	100	117	0	11	4
wearing	**04**	0	0	7	100	17	0	0
leather	**05**	0	0	14	17	100	0	0
wood products	**06**	0	0	2	0	0	100	2
paper	**07**	0	4	49	100	0	0	100

Source: Computation of OST

Table 5: Share of technological fields within industrial sectors with reference to the diagonal elements (horizontal comparison in %), (extraction for the first seven NACE sectors)

industrial sectors		technological fields						
		food **01**	**tobacco** **02**	**textiles** **03**	**wearing** **04**	**leather** **05**	**wood prod-ucts** **06**	**paper** **07**
food	**01**	100	0	1	0	0	0	1
tobacco	**02**	0	100	0	0	0	0	0
textiles	**03**	0	7	100	12	0	4	12
wearing	**04**	0	0	67	100	83	0	0
leather	**05**	0	0	28	3	100	0	0
wood products	**06**	0	0	6	0	0	100	17
paper	**07**	1	1	14	3	0	0	100

Source: Computation of OST

Table 5 shows the contribution of different technological fields within each industrial sector. This perspective refers more closely to the way the database search has been generated. We started from the industrial sectors and looked for their applications in different technological fields. Thus Table 5 shows that firms belonging to sector 4 (wearing) also have technical activities in the technical fields 3 and 5 (textiles, leather). Furthermore, the representation in table 5 is more suitable than the perspective offered in table 4, because the different size of the industrial sectors has an important impact on the outcome with reference to technical fields. For instance, we have to take into account the fact that the pharmaceutical sector, in terms of turnover, employees etc., is much larger than the sector of "other chemicals" (special chemistry) and therefore may dominate the absolute number patents in the technical field of other chemicals, due to technical interconnections of basic and special chemistry.

In the context of this paper, it is not possible to discuss the entire details of the complete concordance matrix. Instead we concentrate on three sectors which are important both in terms of production and also in terms of patent applications, namely Chemicals, Machinery, and Electrical equipment. To study the interconnections between the subsectors in each of these major sectors, we disaggregated them down to 3-digit NACE codes (cf. annex 1).

Within the Chemical industry, the patent activities of Basic chemicals and Pharmaceuticals are much larger than those of the other 4 subsectors (Pesticides, Paints, Soaps and detergents, Other chemicals). The analysis of the distribution of the patent applications of the subsectors by technological fields in Table 6 shows a close technological interconnection between Basic chemicals and Pharmaceuticals due to a mutual transfer of products, and reliance on Organic chemistry. The 4 smaller subsectors have strong linkages to these two large areas, to the extent that the number of patents in these fields are higher than those in their "core" fields. For instance, the Soaps and detergents sector has a high level of patents in Pharmaceuticals, as Cosmetics are a main group within the Pharmaceutical IPC subclass (A61K). The same applies to Other chemicals which comprise preparations for Dentistry and which is a main group of the Pharmaceutical IPC subclass. As the distinction between the chemical subsectors are not very clear-cut, there are good reasons to divide the Chemical industry into two fields only, namely Chemicals and Pharmaceuticals. Decisions on such changes of field delineation will be taken in later stages of the research project together with other amendments, for instance, drop of very small sectors because of statistical instability. But these changes will be recommendations to the user of the concordance who may still apply the full disaggregation of fields.

Table 6: Share of technological fields within industrial subsectors with reference to the diagonal elements (horizontal comparison in %) for the Chemical industry

industrial sectors		technological fields					
		basic chemicals	**pesticides**	**paints**	**pharmaceuticals**	**soaps, detergents**	**other chemicals**
		10	**11**	**12**	**13**	**14**	**15**
basic chemicals	**10**	100	4	3	25	2	4
pesticides	**11**	47	100	2	160	0	0
paints	**12**	166	31	100	29	10	15
pharmaceuticals	**13**	32	5	0	100	1	1
soaps, detergents	**14**	37	6	1	202	100	5
other chemicals	**15**	331	15	26	176	145	100

Source: Computation of OST

In contrast, the subsectors of the mechanical industry prove to be quite independent of each other (table 7). This is surprising, because the subsector of Energy machinery comprises basic mechanical elements that can be used in all other subsectors of the mechanical industry. Within the mechanical subsectors, Agricultural machinery is the smallest in terms of volume of patenting and Special machinery the largest.

Table 7: Share of technological fields within industrial subsectors with reference to the diagonal elements (horizontal comparison in %) for the Machinery industry

industrial sectors		technological fields				
		energy machinery	non-specific machinery	agro machinery	machine tools	special machinery
		21	22	23	24	25
energy machinery	**21**	100	37	0	18	24
non-specific machinery	**22**	55	100	0	6	21
agro machinery	**23**	9	2	100	0	5
machine tools	**24**	3	5	0	100	29
special machinery	**25**	6	8	0	8	100

Source: Computation of OST

In the Electrical industry, the technical interconnection of its subsectors is clearly visible, but less distinct than in the case of the Chemical industry (Table 8). An example is the linkage between Computers and Electronic components, Telecommunications, and Television, reflecting the convergence between information technology and consumer electronics. In absolute numbers of patent applications, the subsectors of Computers, Electronic components, and Telecommunications distinctly dominate, so that it may be useful to put these "traditional" subsectors together.

Table 8: Share of technological fields within industrial subsectors with reference to the diagonal elements (horizontal comparison in %) for the Electrical industry

industrial sectors		computers 28	electric motors o 29	electric distribution 30	accumulators 31	lightening 32	other eltc. equipmt. 33	eltc. components 34	telecommunications 35	television 36
computers	**28**	100	2	5	1	0	6	33	46	15
eltc. motors	**29**	23	100	60	0	0	60	13	19	1
eltc. distribution	**30**	16	18	100	1	1	26	33	26	2
accumulators	**31**	1	1	1	100	5	3	1	6	0
lightening	**32**	46	26	258	3	100	42	131	29	0
other eltc. equipmt.	**33**	235	49	267	0	0	100	251	207	43
eltc. components	**34**	105	4	20	3	0	7	100	71	18
telecommuni-cations	**35**	30	1	3	6	0	3	14	100	18
television	**36**	128	1	2	4	0	4	18	57	100

Source: Computation of OST

The possibility of analysing industrial structures is an interesting side effect of the concordance, but the major aim is, of course, the transformation of technology fields into industrial sectors. This task can be realised by a matrix that is equivalent to Table 4 where the contribution of the industrial sectors to the technological fields is considered. The elements of the transformation matrix do not refer to the diagonal elements, but to the sum of the columns. This approach is illustrated in table 9 using the example of Special machinery. The results of the empirical analysis show that only 41 percent of the patent applications in this technical field come from the Special machinery industrial sector. Although the contributions of all the other sectors are quite small if considered individually, they sum up to nearly 60 percent. This structure can be explained by the definition of Special machinery which involves the production of machines for specific sectors such as the Food, Textiles, Wearing, Paper etc. All these sectors use Special machines and contribute themselves to the technical improvement of these machines.

Table 9: Distribution of the technical field of special machinery on different industrial sectors

sector	field no	%	sector	field no	%
food	01	2	agro-machinery	23	0
tobacco	02	0	machine-tools	24	2
textiles	03	1	special machinery	25	41
wearing	04	0	weapons	26	1
leather	05	0	domestic appliances	27	0
wood products	06	0	computers	28	6
paper	07	4	electric motors	29	0
publishing	08	0	electric distribution	30	0
petroleum	09	3	accumulators	31	0
basic chemicals	10	3	lightening	32	0
pesticides, agro-chemicals	11	0	other electrical equip-ment	33	1
paints varnishes	12	1	electronic components	34	1
pharmaceuticals	13	2	telecommunications	35	1
soaps, detergents	14	3	television	36	0
other chemicals	15	3	medical equipment	37	1
man-made fibres	16	1	measuring instruments	38	1
plastic products	17	2	industrial control	39	0
mineral products	18	1	optics	40	4
basic metals	19	4	watches	41	0
metal products	20	2	motor vehicles	42	3
energy machinery	21	2	other transport	43	1
non-specific machinery	22	2	consumer goods	44	0

Source: Computations of OST and Fraunhofer ISI

The analysis of table 9 illustrates the orientation of the correspondence suggested in this paper. It is linked neither to the sectors of manufacture nor to the sectors of use, but it looks at the sectors where new technological concepts are generated. This approach is based on the thesis that the patents from the different sectors of use, e. g., Paper with reference to Special machinery, cannot be simply added to the activity of the sectors of manufacture. To a certain extent, the inventions of the sectors of use (e.g., Paper) contribute to innovation in the sectors of manufacture (e.g., producer of Paper machines), but they also improve the competitiveness in the sectors of use (e.g., Paper), otherwise the patent applications of firms of the sectors of use would not be rational (e.g., patents on Paper machines by firms of Paper production).

6. Statistical verification of the concordance

This section addresses two main questions:

1. Are there differences, if the concordance is applied to different countries, sizes of firms and to different points in time?
2. How far do the resulting patent data by industry correlate with economic variables such as production, value added, exports and R&D expenditures?

The results presented here are preliminary, as the process of constructing the concordances has just been completed. In particular they partially address the first question only. By the time of the conference there will be a more complete set of results.

As described in section 3 above, we have constructed the concordance on the basis of a sample of firms from different countries. Furthermore we have information about the size of firm (in terms of patent applications) and we also have information on applicants at the beginning and at the end of the 1990s. Thus we are able to apply the concordance (on a micro level) for each of the following categories of firms:

- Large firms in the late 1990s
- Large firms in the early 1990s
- Small firms in the late 1990s
- All firms patenting from France (i.e. with applicant addresses in France)
- All firms patenting from Germany (i.e. with applicant addresses in Germany)
- All firms patenting from the UK (i.e. with applicant addresses in the UK)
- All firms patenting from the US (i.e. with applicant adresses in the US)

In principle this allows us to examine whether differences across countries, size of firm and over time matter for the application of the concordance. At the time of writing we have only been able to analyse the issue of differences across size and over time. Before devising more sophisticated tests, we begin by addressing this issue at a very basic level by looking for some simple correlations. In particular the following questions have been addressed:

- What are the similarities in the distribution of a technology across industries over time (or by size class)?
- What are the similarities in the distribution of technologies within an industry over time (or by size class)?

Table 10 addresses the first of these. As an example, the first row of the table shows that the distribution of Food and beverages technologies across the 44 industrial sectors in the early 1990s is very highly correlated with the same distribution in the late 1990s.

Table 10: Similarities in the distribution of each technical field across industrial classes, according to size and over time

Field	Field Name	Time	Size
1	Food, beverages	0.97	0.97
2	Tobacco products	0.95	0.92
3	Textiles	0.83	0.34
4	Wearing apparel	0.34	0.22
5	Leather articles	0.86	0.79
6	Wood products	0.75	0.67
7	Paper	0.81	0.81
8	Publishing, printing	na	na
9	Petroleum products, nuclear fuel	0.96	0.92
10	Basic chemical	1.00	0.80
11	Pesticides, agro-chemical products	0.95	0.60
12	Paints, varnishes	0.97	0.81
13	Pharmaceuticals	0.99	0.99
14	Soaps, detergents, toilet preparations	0.90	0.34
15	Other chemicals	0.97	0.22
16	Man-made fibres	0.74	0.49
17	Rubber and plastics products	0.95	0.91
18	Non-metallic mineral products	0.94	0.91
19	Basic metals	0.98	0.94
20	Fabricated metal products	0.95	0.94
21	Energy machinery	0.97	0.50
22	Non-specific purpose machinery	0.86	0.74
23	Agricultural and forestry machinery	0.99	0.99
24	Machine-tools	0.95	0.81
25	Special purpose machinery	1.00	0.97
26	Weapons and ammunition	0.85	0.74
27	Domestic appliances	0.95	0.76
28	Office machinery and computers	0.99	0.84
29	Electric motors, generators, transformers	0.78	0.40
30	Electric distribution, control, wire, cable	0.80	0.51
31	Accumulators, battery	0.93	0.36
32	Lighting equipment	0.76	0.86
33	Other electrical equipment	0.72	0.19
34	Electronic components	0.94	0.37
35	Signal transmission, telecommunications	0.91	0.83
36	Television and radio receivers, audiovisual electronics	0.98	0.86
37	Medical equipment	0.96	0.97
38	Measuring instruments	0.89	0.63
39	Industrial process control equipment	0.85	0.28
40	Optical instruments	0.97	0.65
41	Watches, clocks	0.99	0.58
42	Motor vehicles	0.99	0.98
43	Other transport equipment	0.99	0.95
44	Furniture, consumer goods	0.95	0.98

Source: Computations of SPRU

Table 11: Similarities in the distribution of technical fields across each industrial class, according to size and over time

Field	Field Name	Time	Size
1	Food, beverages	0.98	0.86
2	Tobacco products	0.90	0.86
3	Textiles	0.87	0.63
4	Wearing apparel	0.53	0.04
5	Leather articles	0.83	0.48
6	Wood products	0.70	0.83
7	Paper	0.75	0.88
8	Publishing, printing	0.97	0.95
9	Petroleum products, nuclear fuel	0.99	0.82
10	Basic chemical	1.00	0.99
11	Pesticides, agro-chemical products	0.98	0.44
12	Paints, varnishes	0.97	0.88
13	Pharmaceuticals	0.99	0.96
14	Soaps, detergents, toilet preparations	0.96	0.94
15	Other chemicals	0.97	0.82
16	Man-made fibres	0.98	0.24
17	Rubber and plastics products	0.98	0.97
18	Non-metallic mineral products	0.98	0.96
19	Basic metals	0.96	0.90
20	Fabricated metal products	0.94	0.90
21	Energy machinery	0.97	0.74
22	Non-specific purpose machinery	0.97	0.87
23	Agricultural and forestry machinery	1.00	1.00
24	Machine-tools	0.99	0.97
25	Special purpose machinery	0.99	0.96
26	Weapons and ammunition	0.96	0.94
27	Domestic appliances	0.99	0.99
28	Office machinery and computers	0.98	0.91
29	Electric motors, generators, transformers	0.76	0.80
30	Electric distribution, control, wire, cable	0.82	0.74
31	Accumulators, battery	0.99	0.97
32	Lighting equipment	0.89	0.57
33	Other electrical equipment	0.85	0.64
34	Electronic components	0.98	0.85
35	Signal transmission, telecommunications	0.90	0.96
36	Television and radio receivers, audiovisual electronics	0.99	0.71
37	Medical equipment	1.00	1.00
38	Measuring instruments	0.90	0.90
39	Industrial process control equipment	0.89	0.79
40	Optical instruments	0.93	0.54
41	Watches, clocks	0.94	0.59
42	Motor vehicles	1.00	0.97
43	Other transport equipment	0.81	0.92
44	Furniture, consumer goods	0.96	0.94

Source: Computations of SPRU

At the same time similar distributions by size classes are also highly correlated. In general the table shows greater similarities over time than across size classes. Thus for 37 out of 43 fields, the spread across industrial sectors is very similar (all correlations below 0.8 have been defined as being low and are marked in bold (not statistically significantly different from 0 at the 5% level) or in italics (all others below 0.8)). On the other hand there are many more cases of dissimilarity by size: 21 out of 43 correlations are less than 0.8, and in 6 of these they are not statistically different from zero.

Table 11 considers the second question, regarding the similarities in the distribution of technologies within an industry over time (or by size class). The first row of the table shows that the contribution made by each of the 44 areas of technology to the Food industry in the early 1990s is very similar to that made by the same technologies in the late 1990s. It shows that this similarity in the distributions also holds across the size classes in the same industry. The general point made above about greater similarities over time than across size classes applies here. There are only 4 industries where the contribution of different technologies has changed over time. However there are 13 such cases when comparing size classes.

The information in Tables 10 and 11 can be used to identify some of the most problematic areas in the concordances. From Table 10 it is clear that these are Other electrical equipment, Wearing apparel, Industrial process control equipment, Other chemicals, Textiles. Table 11 shows that Wearing apparel, Man-made fibres, Pesticides and Agrochemical products are also problematic. The main reason for instability are low absolute numbers of patent applications or less clearly defined categories (Other electrical equipment, Other chemicals). One solution for improving the concordance could be to aggregate these upto a higher level.

The final question that is pertinent to the evaluation of the concordances is the extent to which there are similarities over time and by size class in the importance of the "core" sector in each area of technology[10] (see the discussion above in Section 5). Table 12 addresses this question over time. It shows for example in the first row that in technologies related to Soaps, Detergents, Toilet preparations the importance of firms whose principle industrial activity is also in this area increased markedly from 37.2% of all patenting in those technologies in the early 1990s to 67.2% in the late 1990s. At the other end of the spectrum, in technologies related to Leather articles, the importance of firms involved in producing Leather articles declined sharply. The main point to note from this table is that there are some important changes in the structure of the concordance over time.

[10] These are the diagonal elements of the matrix described above in section 5.

Table 12: Similarities in the importance of the 'core' sector in each technical field over time

Field	FieldName	Early 1990s	Late 1990s	Diff
14	Soaps, detergents, toilet preparations	37.2	64.6	27.4
16	Man-made fibres	16.1	36.7	20.6
35	Signal transmission, telecommunications	22.4	39.5	17.2
1	Food, beverages	47.6	63.6	16.0
26	Weapons and ammunition	30.4	46.2	15.9
44	Furniture, consumer goods	36.7	51.3	14.6
37	Medical equipment	41.3	54.2	12.8
34	Electronic components	14.0	24.4	10.4
27	Domestic appliances	30.5	39.1	8.6
4	Wearing apparel	0.0	8.3	8.3
18	Non-metallic mineral products	19.0	26.7	7.7
39	Industrial process control equipment	6.3	13.6	7.4
42	Motor vehicles	54.5	61.4	6.9
32	Lighting equipment	29.4	35.5	6.1
41	Watches, clocks	55.6	61.2	5.6
15	Other chemicals	8.2	13.3	5.0
22	Non-specific purpose machinery	11.6	16.4	4.8
30	Electric distribution, control, wire, cable	8.2	12.6	4.3
24	Machine-tools	16.5	20.8	4.3
36	Television and radio receivers, audiovisual electronics	20.2	23.6	3.4
40	Optical instruments	25.4	28.5	3.1
38	Measuring instruments	10.2	13.2	3.1
13	Pharmaceuticals	64.5	65.9	1.3
3	Textiles	10.8	11.6	0.7
17	Rubber and plastics products	23.7	24.4	0.7
25	Special purpose machinery	40.7	41.3	0.6
9	Petroleum products, nuclear fuel	45.6	45.7	0.1
43	Other transport equipment	60.2	60.2	0.0
8	Publishing, printing	0.0	0.0	0.0
10	Basic chemical	35.6	34.4	-1.2
19	Basic metals	38.7	37.4	-1.2
29	Electric motors, generators, transformers	9.6	8.0	-1.6
12	Paints, varnishes	16.7	14.4	-2.3
21	Energy machinery	15.8	13.5	-2.3
11	Pesticides, agro-chemical products	13.6	10.0	-3.6
7	Paper	24.5	20.9	-3.7
28	Office machinery and computers	52.2	48.5	-3.7
31	Accumulators, battery	19.6	15.7	-3.9
20	Fabricated metal products	35.1	31.2	-3.9
23	Agricultural and forestry machinery	75.3	66.9	-8.5
6	Wood products	32.4	23.4	-9.0
33	Other electrical equipment	16.3	3.6	-12.6
2	Tobacco products	72.4	56.1	-16.3
5	Leather articles	44.8	24.6	-20.3

Source: Computations of SPRU

Table 13: Similarities in the importance of the 'core' sector in each technical field by size class

Field	Field Name	Small	Large	Diff
14	Soaps, detergents, toilet preparations	20.0	64.6	44.6
41	Watches, clocks	33.3	61.2	27.9
28	Office machinery and computers	25.5	48.5	23.1
27	Domestic appliances	23.7	39.1	15.3
42	Motor vehicles	47.6	61.4	13.8
40	Optical instruments	16.8	28.5	11.7
11	Pesticides, agro-chemical products	0.0	10.0	10.0
16	Man-made fibres	27.3	36.7	9.4
6	Wood products	14.3	23.4	9.1
43	Other transport equipment	52.5	60.2	7.8
37	Medical equipment	48.3	54.2	5.8
39	Industrial process control equipment	8.6	13.6	5.0
34	Electronic components	19.4	24.4	5.0
26	Weapons and ammunition	43.6	46.2	2.6
23	Agricultural and forestry machinery	64.5	66.9	2.4
13	Pharmaceuticals	64.7	65.9	1.1
36	Television and radio receivers, audiovisual electronics	22.6	23.6	1.0
8	Publishing, printing	0.0	0.0	0.0
12	Paints, varnishes	16.1	14.4	-1.6
44	Furniture, consumer goods	53.2	51.3	-1.9
19	Basic metals	39.8	37.4	-2.4
9	Petroleum products, nuclear fuel	50.0	45.7	-4.3
1	Food, beverages	69.3	63.6	-5.7
7	Paper	27.1	20.9	-6.2
25	Special purpose machinery	48.9	41.3	-7.6
3	Textiles	21.4	11.6	-9.9
17	Rubber and plastics products	35.5	24.4	-11.1
33	Other electrical equipment	15.1	3.6	-11.5
24	Machine-tools	33.8	20.8	-13.0
15	Other chemicals	27.5	13.3	-14.2
20	Fabricated metal products	46.8	31.2	-15.6
35	Signal transmission, telecommunications	55.6	39.5	-16.1
22	Non-specific purpose machinery	32.7	16.4	-16.3
18	Non-metallic mineral products	45.5	26.7	-18.8
38	Measuring instruments	35.0	13.2	-21.8
4	Wearing apparel	32.0	8.3	-23.7
21	Energy machinery	37.7	13.5	-24.2
2	Tobacco products	81.8	56.1	-25.7
32	Lighting equipment	61.5	35.5	-26.1
5	Leather articles	52.6	24.6	-28.1
31	Accumulators, battery	44.4	15.7	-28.8
10	Basic chemical	65.0	34.4	-30.5
29	Electric motors, generators, transformers	38.8	8.0	-30.8
30	Electric distribution, control, wire, cable	44.3	12.6	-31.8

Source: Computations of SPRU

Finally Table 13 examines the similarities in the importance of the 'core' sector in each technical field by size class. A casual glance at this table shows that in a large number of technical fields there are major important differences across large and small firms in the importance of the 'core' sector. In 25 out of the 43 cases there is a difference in the share of more than 10%. Again this suggests that the concordance based on large firms would in general be different from the concordance based on small firms.

The next steps for the evaluation of the concordance are comparisons over different countries, as mentioned above, including applicants from France, United Kingdom, Germany and the United States. Furthermore, a shift from the more descriptive evaluation done so far to a multivariate approach will be done, taking into account all dimensions of the concordance (technologies, sectors, firm size or time respectively) simultaneously.

After the checks for 'reliability' of the concordance, its 'validity' will be evaluated using different indicators such as value added, production, R&D-expenditure or foreign trade. The concordance for linking technologies and sectors has to prove its explanatory power both in simple and more sophisticated economic models.

One further point has to be evaluated in detail in subsequent research. For the construction of the concordance firms classified as 'holdings' have been dropped, because the 'core' activity (in a technological and sector-based sense) could not be identified. In some cases such 'holdings' stand for a relevant amount of patents, which has not been taken into account so far. The impact of this exclusion has to be clarified.

7. Conclusions

The empirical analyses underlying the building of the concordance shows that a straightforward definition of industrial sectors by technologies is not appropriate. The two main reasons are that sometimes there is a strong technological interconnection between different sectors, and secondly that large firms produce a broad spectrum of technologies. The suggested concordance can be used for international comparisons, as it refers to international classifications, namely NACE and ISIC for industrial sectors and IPC for patents. With 44 sector fields, the concordance has a reasonable level of disaggregation. A further differentiation would not be useful, as the economic data for international comparisons are not available in a finer breakdown, and the technical interconnections between the subsectors would become too strong. Higher aggregation levels can be achieved by a mere combination of subsectors.

The correspondence has a sound empirical basis, as it does not entirely rely on expert assessment in a technological perspective, but on the patent activities of industrial sectors, determined by a very large sample of enterprises. Moreover, the application of the concordance to specific examples requires a limited amount of work. Database searches have to be performed for only 44 technological fields, defined by a set of IPC subclasses, whereby the results can be transformed using a 44x44 matrix into industrial sectors. Therefore the searches do not require in-house databases, but can be realised by online databases, too. The transformation does not need special software developments and can be done by standard calculation programs.[11]

A specific advantage of the correspondence is the possibility of analysing industrial structures, for instance, by making comparisons across countries, looking for changes over time, or examining differences between large and small enterprises. For such purposes, the technical definitions are kept invariant, whereas different data sets are used for the empirical construction of structural matrices.

Further research needs to be conducted on the validity of the concordance, in particular in view of international comparisons and comparisons over time. For instance, the necessary intervals of updates have to be checked, linked to structural changes in technology and industry. In addition, the outcome of the suggested correspondence has to be compared to that of other concordances in order to assess its reliability.

---

[11] Search instructions for the 44 IPC-based fields and an Excel-sheet for the computations will soon be available for download on the web-pages of Fraunhofer-ISI (www. isi.fhg.de).

References

Evenson, R./ Putnam, J. (1988): The Yale-Canada patent flow concordance, Yale University, Economic Growth Centre Working Paper.

Johnson, D.K.N. (2002): The OECD Technology Concordance (OTC): Patents by Industry of Manufacture and Sector of Use. STI Working Papers 2002/5. Paris: OECD.

Verspagen, B./ Morgastel, T. v./ Slabbers, M. (1994): MERIT concordance table: IPC – ISIC (rev. 2), Maastricht: MERIT Research Memorandum 2/94-004.

Blind, K./ Anderson, B./ Evangelista, R./ Hipp, C. (2003): Patents in the service industries. Report to the European Commission, Karlsruhe: Fraunhofer ISI, forthcoming.

Schmoch, U./Grupp, H./Mannsbart, W./Schwitalla, B. (1988): Technikprognosen mit Patentindikatoren, Köln: Verlag TÜV Rheinland.

Annex 1: Definition of 44 sectoral fields by NACE codes

Field no	NACE	Description
1	15	Food, beverages
2	16	Tobacco products
3	17	Textiles
4	18	Wearing apparel
5	19	Leather articles
6	20	Wood products
7	21	Paper
8	22	Publishing, printing
9	23	Petroleum products, nuclear fuel
10	24.1	Basic chemical
11	24.2	Pesticides, agro-chemical products
12	24.3	Paints, varnishes
13	24.4	Pharmaceuticals
14	24.5	Soaps, detergents, toilet preparations
15	24.6	Other chemicals
16	24.7	Man-made fibres
17	25	Rubber and plastics products
18	26	Non-metallic mineral products
19	27	Basic metals
20	28	Fabricated metal products
21	29.1	Energy machinery
22	29.2	Non-specific purpose machinery
23	29.3	Agricultural and forestry machinery
24	29.4	Machine-tools
25	29.5	Special purpose machinery
26	29.6	Weapons and ammunition
27	29.7	Domestic appliances
28	30	Office machinery and computers
29	31.1	Electric motors, generators, transformers
30	31.2, 31.3	Electric distribution, control, wire, cable
31	31.4	Accumulators, battery
32	31.5	Lightening equipment
33	31.6	Other electrical equipment
34	32.1	Electronic components
35	32.2	Signal transmission, telecommunications
36	32.3	Television and radio receivers, audiovisual electronics
37	33.1	Medical equipment
38	33.2	Measuring instruments
39	33.3	Industrial process control equipment
40	33.4	Optical instruments
41	33.5	Watches, clocks
42	34	Motor vehicles
43	35	Other transport equipment
44	36	Furniture, consumer goods