DISCUSSION PAPER SERIES

No. 5971

CLAIMING MORE: THE INCREASED VOLUMINOSITY OF PATENT APPLICATIONS AND ITS DETERMINANTS

Nicolas van Zeebroeck, Bruno van Pottelsberghe and Dominique Guellec

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP5971.asp

ISSN 0265-8003

CLAIMING MORE: THE INCREASED VOLUMINOSITY OF PATENT APPLICATIONS AND ITS DETERMINANTS

Nicolas van Zeebroeck, Universite Libre de Bruxelles
Bruno van Pottelsberghe, European Patent Office, Université Libre de Bruxelles and CEPR
Dominique Guellec, OECD

Discussion Paper No. 5971
December 2006

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 5971

December 2006

ABSTRACT

Claiming More: The Increased Voluminosity of Patent Applications and its Determinants*

The joint increase in the number and size of patents filed around the world puts the patent system under pressure. This paper analyses the sources of this surge in number of claims and pages of patent applications at the EPO. Four hypotheses are scrutinized: the diffusion of national drafting practices, the increasing complexity of inventions, the emergence of new sectors, and new patenting strategies. The results show that the increasing voluminosity is explained by all these hypotheses and suggest that the diffusion of the US model through the PCT is one of the major factors driving the size of EPO patent applications.

JEL Classification: O31, O34 and O50
Keywords: claim drafting, IP strategy, patent applications, patent systems and patent voluminosity

Dominique Guellec
OECD
Rue André Pascal, 2
75775 Paris
France
Email: dominique.guellec@oecd.org

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=150589

Bruno van Pottelsberghe
ULB
Solvay Business School
Centre Emile Bernheim CP145/01
Av. Roosevelt 21 1050 Brussels
Belgium
Email: bruno.vanpottelsberghe@ulb.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=150075

Nicolas van Zeebroeck
ULB
Solvay Business School
Centre Emile Bernheim CP145/01
Av. Roosevelt 21 1050 Brussels
Belgium
Email: nicholas.vanzeebroeck@ulb.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165661

* We gratefully acknowledge the helpful comments and suggestions we received from Eugenio Archontopoulos, Wolfram Förster and Niels Stevensborg of the EPO, Robert M. Hunt of the Federal Reserve Bank of Philadelphia and from participants of the Fifth EPIP conference in Copenhagen in March 2005, of the Freeman Centre Seminar Series 2004/05 held at SPRU (University of Sussex), of the Second ZEW Conference on "Economics of Innovation and Patenting" in Mannheim in September 2005 and of the 26th AEA Conference on "Innovations & Intellectual Property" in Paris in October 2005.

Submitted 01 November 2006

1. Introduction

Patents have a strategic importance in knowledge economies, and they have increasingly become an object of intense covetousness and disputes between industry players. At the very heart of this rivalry are the patented claims, which define the legal scope of the invention for which protection is being granted. Consequently, it is often in the interests of patentees to obtain the broadest possible protection by claiming more in their patent applications, which may result into larger documents.[1]

In December 2000, in an article titled "*Patently Ridiculous*", New Scientist evoked a drastic increase in the length of patent applications over the world. The author attributed this surge mainly to a stampede to file patents on biotech inventions and he accused it of bringing the patent system to its knees.[2] Once something totally unconceivable, applications of over thousand pages are now frequently filed at the European Patent Office (EPO)[3] and other patent offices around the world, and several applications even reached 100,000 pages or up to 20,000 claims in recent years. This effect, combined with another drastic surge in the number of applications themselves, is in fact generating such a workload on patent offices that the EPO has to deal with a huge backlog of applications still to be searched for or examined, as pointed out by Archontopoulos *et al.* (2006).[4]

To a larger scale, the EPO has witnessed a radical surge in the size of patent applications over the past two decades no matter it is measured in terms of number of claims or pages in the documents. While the former indicator has almost doubled from 10 to about 18 claims per application between 1980 and 2002, the latter has witnessed a similar evolution in an even shorter period, from 14 pages in 1988 to 30 pages per application in 2002. This evolution has resulted in a heavier workload for the EPO (as well as other patent offices in the world), as a larger application takes more time to process. Efficiency and quality of patent processing activities have then been at the centre of public debates.

To explain this evolution in patent voluminosity, there surely is a general trend toward more complexity and more information associated with new technologies and cumulative inventions (e.g. the user manuals of home electronic devices are much larger and complex nowadays than 15 years ago). But in addition to this trend, anecdotal evidence, highlighted by Archontopoulos *et al.* (2006), suggest that the international diffusion of the US model through the Patent Cooperation Treaty (PCT)[5] as well as the development of emerging technological areas such as biotechnologies are at

---

[1] In what follows, the term "applicant" will be used to refer to the person, institution or firm having applied for the patent and who is, under the European patent system, the initial owner of the patent rights (the equivalent of the "assignee" in the US). By virtue of this "first to file" principle, the term "patentee" may refer either to the applicant/assignee or to the actual inventor of the patent.

[2] Fox B. and A. Coghlan, "Patently Ridiculous" in New Scientist, December 2000.

[3] Established by the Convention on the Grant of European Patents (EPC) signed in Munich 1973, the EPO is the outcome of the European countries' collective political determination to establish a uniform patent system in Europe. The EPO was set up by the contracting states to the EPC with the aim of strengthening co-operation between the countries of Europe in the protection of inventions. This was achieved by adopting the EPC, which makes it possible to obtain such protection in several or all of the contracting states by a single patent grant procedure, and establishes standard rules governing the treatment of patents granted by this procedure. By filing a single application in one of the three official languages (English, French and German) it is possible to obtain patent protection in some or all of the EPC contracting states. The resulting patent is, however, not a single patent but rather a bundle of national patents.

[4] See chapter 5 of Jaffe and Lerner (2004) for an illustration of this issue at the USPTO.

[5] The Patent Cooperation Treaty (PCT), subscribed to by most of the members of the Paris Convention and supervised by the World Intellectual Property Organization (WIPO), offers inventors a major way of deferring patenting expenses. Under this Treaty, one can file an "International Patent Application", which doesn't turn into some sort of international patent, but primarily acts as a vehicle to buy a period of time within which to proceed with national or regional (such as the EPO) patent applications. Instead of having only a twelve month time period within which an inventor must file

the origin of this obvious race to larger files. If these factors most probably play a significant role, an integrated and comprehensive view has to be implemented in order to understand the sources of this escalation.

This is exactly the objective of this paper: to provide a coherent analysis of the determinants of patent applications' voluminosity and to identify the major drivers of its surge, which requires exploring the very intimate anatomy of patent applications. The approach relies on a quantitative model applied to a unique database with data on more than one and a half million EPO applications, filed between 1982 and 2004. Although the number of claims has frequently been considered in the literature as a determinant or indicator of patent value (e.g. Lanjouw and Schankerman (1997, 1999), Reitzig (2004)), the number of pages, to the best of our knowledge, has not yet been scrutinized, at least in such a systematic way. Moreover, the determinants of neither of these indicators have been subject to much investigation so far.

The paper is organized as follows. Section 2 reviews possible interpretations of the number of patent claims and pages and elaborates on the reasons why voluminosity matters in the patent system. Section 3 is devoted to the development of a set of four broad hypotheses associated with an application's characteristics, which could contribute to determine its size and to explain the observed phenomena: *i)* national practices and the international diffusion of the US model, *ii)* the technical complexity of inventions, *iii)* the impact of emerging technological sectors in the patent field and *iv)* patenting strategies. The econometric model, its results and interpretations of the determinants of patent voluminosity follow in section 4. Concluding remarks are exposed in section 5.

The results show that the voluminosity of patent applications at the EPO, as measured with the number of claims or the number of pages, can be explained by the four broad hypotheses to different extents. First of all, the international diffusion of national models and regulations affects EPO filings *via* the PCT option, especially from the USA. This phenomenon is reinforced by the fast emergence of certain sectors, namely biotechnologies, telecommunications and computers. In addition, the increasing technological complexity of inventions and to a lower extent some elements of patenting strategies appear as major determinants of the voluminosity. Finally, the numbers of claims and pages seem subject to the same contingencies, notwithstanding a few discrepancies. We suggest that the major differences across the US (and other anglo-saxon countries) on the one hand, and continental Europe on the other hand, illustrate an ultimate influence of the fundamental differences between Common and Civil Law systems and their aftermath on the European patent system.

2. The increased voluminosity of patent applications

A patent document is made of bibliographic data – providing information on the applicant, inventor, technological classes, references to the existing prior art, etc. –, the specification or description of the invention, the claims – what the patentee is claiming exclusive rights on, aka the scope of the patent –, and finally – but optionally – some illustrations supporting the specification and claims in the form of drawings, listings, gene sequences, etc. As they constitute the legal core of a patent, the

---

foreign applications in order to claim priority, with the PCT, the inventor can gain an additional eighteen months before having to incur the relatively large expenses of completing the applications at each of the designated offices. Those additional months can be crucial to the exploitation of an invention. They may give the inventor additional time to raise the funds required to file patent applications in a large number of countries, or provide additional time within which to gauge the economic importance of the invention, or to find licensees or even partners for the project.

claims have been subject to numerous investigations in the economic literature. Their most appropriate interpretation is however still unclear.[6]

The number of claims might reflect a **broader scope of protection** since more subject matter is included. This is the dominant interpretation in the literature (e.g. Tong and Frame, 1994; Lanjouw and Shankerman, 1999), and it is confirmed by practitioners in many cases. Not all. The breath of a patent is often tied to the wording of claims – e.g. replacing the word "rodent" by the word "mouse" will drastically shrink the scope of a patent. Adding claims could even in certain cases signal a narrower filing – e.g. listing three types of rodents takes three claims, while just mentioning rodents takes one. However it is empirically sound to consider that the number of claims is overall positively correlated with the scope of protection.[7]

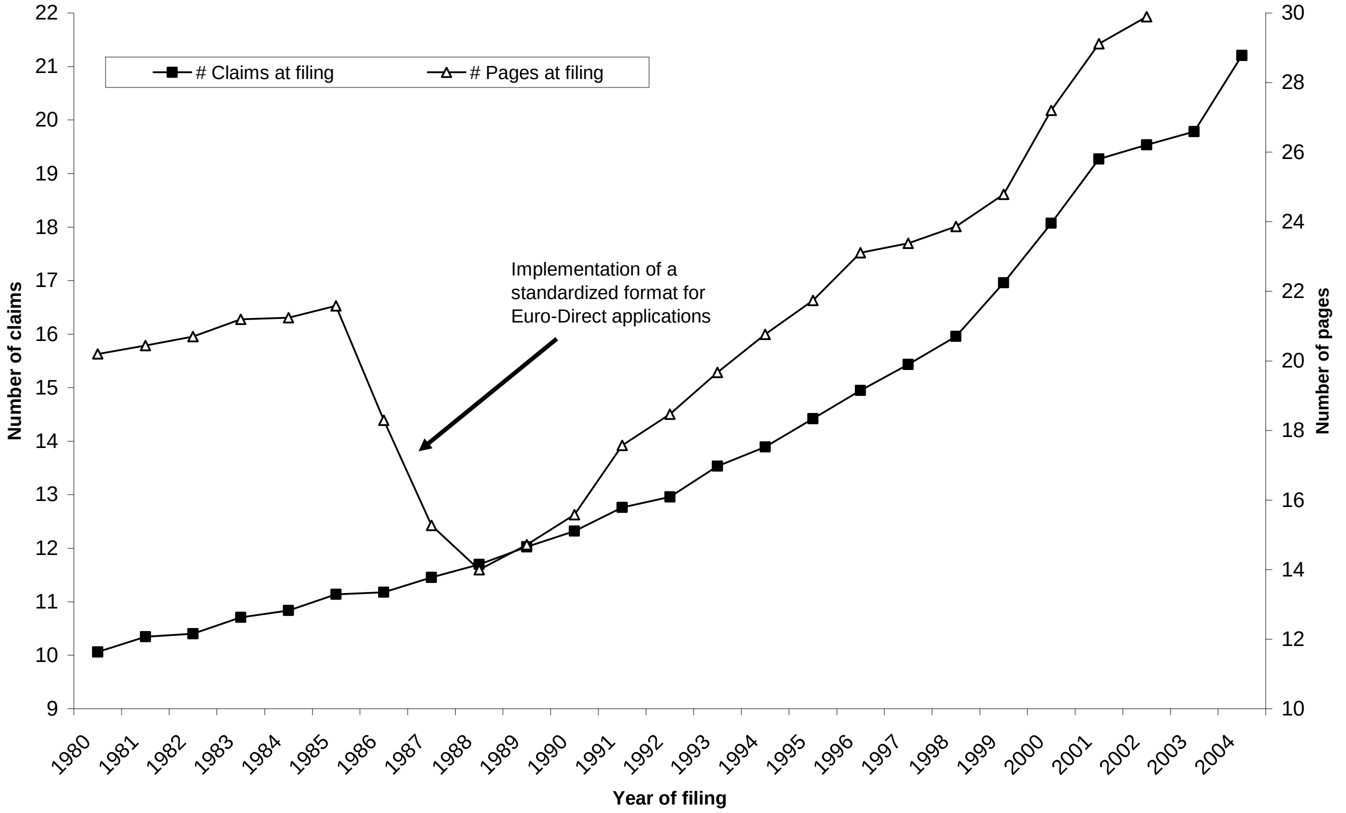
More claims may nevertheless also denote a **more detailed definition** or explanation of the protected area, adding precision: instead of giving a generic term which could be somewhat vague, the applicant will list extensively and individually all components of the subject matter. The purpose could be to secure the legal validity of the patent in case of licensing or in case of litigation. It could also be to construct "fall back positions" in the course of the examination procedure or in view of the application of the doctrine of equivalents. By having a series of claims partly overlapping, partly fitted into each other, the applicant has the possibility of fine tuning the scope of protection in front of the examiner's objections and to maximize her chances to be able to claim infringement under the doctrine of equivalents. This is of particular relevance in the US where something is deemed equivalent only if the variation between the features of the infringing device or process and the patented claim are "insubstantial". The *Festo* case decided by the Supreme Court in 2002 has nevertheless reasserted the "file history estoppel", by which a patentee cannot claim back subject matter that she abandoned (with amendments to her application) in the course of patent prosecution. This principle is a restriction to the use of the doctrine of equivalents provided that the amendments were made with the intent of narrowing the patented claims so as to avoid some problem that would void an overbroad patent, such as overlap with a different patent.[8]

Finally, the number of claims may be the result of some "**strategic choice**", as clearly suggested by Stevnsborg and van Pottelsberghe (2007). More claims may indeed also betray a willingness of the applicant to hide the true invention in the middle of many "non inventions", using vagueness as a weapon. This could be either because the applicant wants to deceive competitors or patent examiners, or because she does not know herself at the time of filing what the real invention is. For instance the inventor of a chemical product, after experiments, would identify a family of compounds, some members of which would have a certain property, but she does not know which one of them has it. She will therefore list all the members in the application, in order not to miss the right one. In order to write such patent applications, it is possible to use some *ad-hoc* software (commercially available) that will combine various works or sequences of letters in as many ways as needed, so that one might draft automatically a patent application with thousands of claims coming from the same mould and covering a broad field, a tiny part of which only is of any relevance.

[6] The distinction between independent and dependent claims might be somewhat helpful in this respect, but such a distinction does unfortunately not exist in EPO data for there should normally be only one independent claim (possibly one per category of invention) in each application.
[7] The number and scope of claims are usually restricted during the examination, so that granted patents are narrower than applications (on average, granted patents have one to two claims less than when applied).
[8] Wikipedia, http://en.wikipedia.org/wiki/Prosecution_history_estoppel.

Such a strategy makes obviously sense in the case of inventions still in the early stage, not mature enough for being completely and precisely described. Hence one would expect the following strategy from the applicant: first, filing a broad application, with many claims most of which are irrelevant, then filing divisional applications, possibly over several generations (i.e. divisionals of divisionals) so as to restrict progressively the scope of protection while research is advancing and still benefiting from the earliest priority. Such behaviour may be suspected for instance in the case of a few gigantic patent applications filed in 2004 at the WIPO by Angiotech International AG, a Swiss biotech company, with a US priority and 5-digit numbers of claims, up to 19,368.

Figure 1 – Average number of claims and pages in incoming applications at EPO (1980-2004)



Source – own calculations based on EPO data

Notably resulting from the number and length of the claims, the number of pages might be seen as the extent of the disclosure of the invention, without presuming of its quality. The more thorough the description and the more exhaustive the drawings, the more pages the application will contain and the more it may divulge on the invention. In parallel with the claims, the average number of pages in patent applications has witnessed a drastic surge over the past two decades with record applications of up to 140,000 pages filed at the WIPO in 2000.[9] Here again, more voluminous files might just be needed for more complex inventions, requiring more wording, details and possibly illustrations to be entirely disclosed. Or it might be the result of a deliberate willingness to contribute to hiding the true invention in the middle of gigantic applications, creating a new variant of submarine patents.[10]

[9] UNDP (2001), p. 103.

[10] "Submarine patent" usually referred to a patent published long after the original application was filed. Like a submarine, it stayed under water (unpublished) for long, then emerged (i.e. has been granted and published), and surprised the whole market. Since applications are now published after 18 months in most countries, no matter they are already granted or not, submarine patents *per se* have disappeared. Jumbo applications are nowadays the new "*de facto*"

An implicit and key question is whether such jumbo-applications do matter. Obviously, they do. As does the overall increasing size of patents, for it has an impact on the patent system, the economy and the society at large. To the economy, more claims often mean that broader protection is sought, inducing a higher cost to society, and possibly more uncertainty for competitors during the examination procedure and beyond. To patent offices, more claims or pages mean more work, hence more resources allocated to searching and examining the files, which induces an additional pressure on quality. What is more, this growing size of applications comes along with a very sharp increase in the number of patent filings themselves, which have been multiplied by 2 over the past 10 years. As a result, the total number of claims and pages to be examined by the EPO for instance is now growing exponentially.

Table 1 - Summary statistics of voluminosity indicators (endogenous variables)

Variable	Period	Obs	Mean	St. Dev.	Min	Max	Mode	Median	Annual Growth
# Claims at filing	1982-2004	1 551 459	14,60	12,08	1	651	10	11	2,5%
# Claims at filing	1988-2002	1 147 567	14,74	12,11	1	592	10	11	3,2%
# Pages at filing	1988-2003	1 136 677	21,45	35,95	1	9786	6	13	5,3%

Source – own calculations based on EPO data

The increase in the average number of claims and pages in applications filed to the EPO is depicted in figure 1 and summary statistics for these variables are provided in table 1. It shows that these numbers have increased over the past two decades, and especially since the nineties with a 50% increase in the number of claims (21 claims per patent on average in 2004 against 14 in 1994) and pages (30 pages per patent in 2002 against 20 in 1994). Other patent offices around the world have experienced a similar phenomenon.

3. Hypotheses on the determinants of patent voluminosity

In order to explain the increasing voluminosity of patent applications at the EPO, four broad candidate hypotheses have been identified. These hypotheses are to be tested with a dataset composed of all patent applications at the EPO filed between 1982 and 2004, which makes 1,551,769 filings.

H1: National practices and the diffusion of the US model

Patent rules and procedures in the US are said to generate more claims than in Europe or Japan.[11] It is due to the American legal tradition, as well as to specific explicit rules (claims cannot be cross-referred in the US as much as they are in Europe, forcing the applicant to add claims instead of strengthening the overall structure of the application), and to the application of the doctrine of equivalents and the file history estoppel, reasserted by the *Festo* decision of the Supreme Court, which influenced patentees' behaviour since the early 2000s.[12] According to this principle, a party

submarines patents. The hidden claims, regularly unknown by the applicant herself at the time of filing, can only be identified by text mining techniques.

[11] See Jensen *et al.* (2005) for a comparison of grant procedures in the three main offices (USPTO, EPO and JPO).

[12] At the EPO, it is required that the "*number of the claims shall be reasonable in consideration of the nature of the invention claimed*." (Rule 29(5) EPC). Under Rule 29(2) EPC, there should be only one single independent claim in each "category" (product, process, apparatus or use). More may be allowed "*where it is not appropriate, having regard*

who makes a change to a patent application to accommodate the requirements of patent law cannot claim indirect infringement of an element that was narrowed by that change.[13] Hence to be able to use the doctrine of equivalents, applicants tend to embed in their draft very detailed fall-back positions that they can use in case of litigation for saving as much as possible of the protection afforded by their patent. One could also argue that patents play a more crucial role in competitive processes in the US than in other parts of the world, so that companies will invest more in patent drafting, prosecuting and enforcement. They would then try to obtain broader protection, as the return on such extension is higher.[14]

Conversely, the Japanese system is known for its low number of claims composing each patent. Therefore, when the protection of a Japanese invention is extended abroad, several priority patents are often merged to form a single US or EPO application.[15] This practice might also result in EPO patent applications of different sizes than the average patent applications from other countries.

The assumption (H1) is that the American model tends to diffuse to the rest of the world. Patents are indeed increasingly applied within a worldwide strategy. Therefore companies draft one single patent – often with a US template – then apply it to all other offices, which is largely encouraged and facilitated by the PCT procedure. By lengthening the international “waiting” phase from one year to 30 months, the PCT option provides more time to the applicants to assess the economic value of their invention before taking the decision to file abroad or not, which induces significant expenses. If the patent is effectively filed abroad after 30 months, the total cost is slightly higher than the non PCT route towards the EPO patent (Dernis *et al.*, 2001). If the patent is dropped into the public domain, the total cost of the patenting process is lower than direct EPO applications. This may explain why the PCT option has met an increasing success over the past 20 years, from 15% of EPO applications in 1985 to 50% in 2000, a figure which is currently stable. Furthermore, the PCT does not involve any additional fee for excess claims and allows (under PCT Rules 40 and 68) to separate groups of claims to be directed to separate inventions in the International Search and Preliminary Exam in case the unity of invention is lacking. Hence, it is all easy and costless for applicants to draft more claims than may be needed when their PCT application is filed to ultimately obtain protection in Europe.

Additionally, different fee regimes in place in the various patent offices around the world may further influence the number of claims and pages. At the EPO, additional fees are incurred when the total number of claims exceeds 10, against 20 in the US system. This difference influences patent drafting styles, as illustrated by Archontopoulos et al. (2006).

To capture these potential factors in the empirical model, a set of dummy variables has been computed. It is described in table 2: *PCT* takes the value 1 if the application was filed under the PCT option and 0 in case of a Euro-Direct filing with or without an earlier national priority; *USPR* identifies those applications filed with a US priority by a non-US applicant, which is the case for a bit more than 3% of EPO applications, as shown in table 2; *USAP* takes the value 1 for applications filed by a US applicant with a non-US priority, representing about 1.5% of applications; and 19

---

to the subject-matter of the application, to cover this subject-matter by a single claim.” To the contrary, a U.S. application may have a multiplicity of independent claims.

[13] Wikipedia, http://en.wikipedia.org/wiki/Festo_Corp._v._Shoketsu_Kinzoku_Kogyo_Kabushiki_Co.

[14] See Hall *et al.* (2004) for an in-depth analysis of US patent litigation processes.

[15] Although the cost of filing a patent application at the EPO will depend – among other factors – on the number of claims it contains, it is usually less costly to file a single application with several claims rather than a number of applications with fewer claims, notwithstanding the unity of the claimed protection. Van Pottelsberghe and François (2006) provide in-depth simulations of the price structure of patent filings.

Table 2 - Summary statistics of exogenous variables (1)

Variable	Type	Obs	Mean	St. Dev.	Min	Max	An. Growth (2)
H1: Internationalisation of application routes and diffusion of US modes (Reference for countries = France)							
PCT Filing	DUM	1 551 769	0,380	0,49	0	1	9,85%
Non US Applicant with US priority	DUM	1 551 769	0,035	0,18	0	1	6,72%
US Applicant with non US priority	DUM	1 551 769	0,016	0,13	0	1	0,19%
AT Applicant	DUM	1 551 769	0,010	0,10	0	1	0,50%
AU Applicant	DUM	1 551 769	0,007	0,08	0	1	1,15%
BE Applicant	DUM	1 551 769	0,009	0,10	0	1	1,62%
CA Applicant	DUM	1 551 769	0,012	0,11	0	1	2,66%
CH Applicant	DUM	1 551 769	0,036	0,19	0	1	-1,60%
DE Applicant	DUM	1 551 769	0,199	0,40	0	1	-0,84%
DK Applicant	DUM	1 551 769	0,007	0,08	0	1	3,17%
ES Applicant	DUM	1 551 769	0,005	0,07	0	1	7,52%
FI Applicant	DUM	1 551 769	0,010	0,10	0	1	7,61%
FR Applicant	DUM	1 551 769	0,075	0,26	0	1	-1,76%
GB Applicant	DUM	1 551 769	0,052	0,22	0	1	-3,43%
IL Applicant	DUM	1 551 769	0,005	0,07	0	1	9,90%
IT Applicant	DUM	1 551 769	0,033	0,18	0	1	1,05%
JP Applicant	DUM	1 551 769	0,179	0,38	0	1	1,33%
KR Applicant	DUM	1 551 769	0,008	0,09	0	1	32,04%
NL Applicant	DUM	1 551 769	0,037	0,19	0	1	0,89%
SE Applicant	DUM	1 551 769	0,021	0,14	0	1	-0,13%
US Applicant	DUM	1 551 769	0,289	0,45	0	1	0,29%
Applicant from the ROW	DUM	1 551 769	0,024	0,15	0	1	3,38%
Filing in a non-EPO (DE/EN/FR) language	DUM	1 551 769	0,077	0,27	0	1	5,25%
H2: Technological complexity							
# Inventors	DISC	1 551 769	2,407	1,77	0	53	1,36%
# IPC-7	DISC	1 534 018	1,921	1,33	1	49	0,40%
# Backward Patent Citations	DISC	1 461 657	4,530	2,89	0	125	0,50%
# Non Patent Citations	DISC	1 461 657	0,953	1,84	0	170	2,74%
H3 Emerging sectors							
JC-01 - Industrial Chemistry	DUM	1 551 769	0,113	0,32	0	1	-1,73%
JC-02 - Organic Chemistry	DUM	1 551 769	0,136	0,34	0	1	-0,69%
JC-03 - Polymers	DUM	1 551 769	0,094	0,29	0	1	-1,43%
JC-04 - Biotechnology	DUM	1 551 769	0,121	0,33	0	1	2,32%
JC-05 - Telecommunications	DUM	1 551 769	0,052	0,22	0	1	7,18%
JC-06 - Audio/Video/Media	DUM	1 551 769	0,049	0,22	0	1	3,21%
JC-07 - Electronics	DUM	1 551 769	0,074	0,26	0	1	-0,16%
JC-08 - Electricity & Elec. Machines	DUM	1 551 769	0,108	0,31	0	1	-0,71%
JC-09 - Computers	DUM	1 551 769	0,048	0,21	0	1	4,61%
JC-10 - Measuring Optics	DUM	1 551 769	0,089	0,29	0	1	-0,32%
JC-11 - Handling & Processing	DUM	1 551 769	0,125	0,33	0	1	-1,35%
JC-12 - Vehicles & Gen. Technology	DUM	1 551 769	0,099	0,30	0	1	0,13%
JC-13 - Civil Engineering / Thermodynamics	DUM	1 551 769	0,089	0,28	0	1	-0,84%
JC-14 - Human Necessities	DUM	1 551 769	0,111	0,31	0	1	0,67%
H4: Patenting strategies							
# Priorities	DISC	1 551 769	1,171	1,29	0	482	0,23%
Application has issued divisionals	DUM	1 543 076	0,020	0,14	0	1	12,14%
Application is a divisional	DUM	1 543 076	0,024	0,15	0	1	13,68%
# Cumulative Filings (5 years)	DISC	1 551 743	420,664	1082,48	0	11 111	4,39%
Occasional (no filing in 4 prev, yrs)	DUM	1 551 743	0,234	0,42	0	1	-1,57%
Year of filing							
1982	DUM	1 551 769	0,018	0,13	0	1	
1983	DUM	1 551 769	0,020	0,14	0	1	
1984	DUM	1 551 769	0,023	0,15	0	1	
1985	DUM	1 551 769	0,025	0,15	0	1	
1986	DUM	1 551 769	0,027	0,16	0	1	
1987	DUM	1 551 769	0,029	0,17	0	1	
1988	DUM	1 551 769	0,034	0,18	0	1	
1989	DUM	1 551 769	0,037	0,19	0	1	
1990	DUM	1 551 769	0,041	0,20	0	1	
1991	DUM	1 551 769	0,038	0,19	0	1	
1992	DUM	1 551 769	0,039	0,19	0	1	
1993	DUM	1 551 769	0,039	0,19	0	1	
1994	DUM	1 551 769	0,040	0,20	0	1	
1995	DUM	1 551 769	0,042	0,20	0	1	
1996	DUM	1 551 769	0,046	0,21	0	1	
1997	DUM	1 551 769	0,051	0,22	0	1	
1998	DUM	1 551 769	0,058	0,23	0	1	
1999	DUM	1 551 769	0,063	0,24	0	1	
2000	DUM	1 551 769	0,069	0,25	0	1	
2001	DUM	1 551 769	0,072	0,26	0	1	
2002	DUM	1 551 769	0,071	0,26	0	1	
2003	DUM	1 551 769	0,071	0,26	0	1	
2004	DUM	1 551 769	0,047	0,21	0	1	

DUM = Dummy variable | DISC = Discrete variable

(1) International PCT filings excluded (2) Average annual growth rates of the share of filings concerned in case of dummy variables

country dummies identify the country of residence of the applicants.[16] Table 2 shows that 29% of EPO applications originate from the USA, 20% from Germany and 18% from Japan. When the number of pages is the dependent variable, the language of the publication must be taken into account with the variable *NO_EPL*. This variable is equal to 1 when the number of pages refers to another language than the three official EPO languages (English, French or German).[17]

H2: Technical complexity

As technology becomes more complex, more words may be required to describe and claim it. Notably because a dwarf standing on the shoulders of a giant may really see farther than a giant himself, architectural inventions tend to lead to more and more complex inventions and technologies over time, produced by larger and larger teams of inventors with complementary skills and expertise. If this assumption had to be true and if more complex technologies do require longer descriptions, one may expect that the rise in technology complexity will drive the voluminosity of subsequent patent applications.

This hypothesis is measured with 4 discrete variables: *INV*, representing the number of inventors listed in the application; *IPC8*, the number of IPC (International Patent Classification) classes at 8 digits associated with the invention, which denotes the technological diversity embodied in the invention, i.e. an "architectural invention" resulting from a process of combination of existing ideas and devices (see for instance Guellec and van Pottelsberghe, 2000, 2002);[18] *BPC*, the number of citations made to previous patents, which indicates a larger use of prior "patented" art; and *NPC*, the number of citations to the scientific literature made by the application, which allows to identify science-based inventions. Table 2 indicates that the average EPO application has been produced by 2.4 inventors (with a maximum of 53), covers about two IPC8 classes and up to 49, and makes 4.5 citations to earlier patents (with a maximum of 125) and only about 1 reference to the scientific literature (up to 170, however).

H3: Emerging sectors

The wording space required for the codification of an invention may vary substantially across technological areas. The vocabulary of more recent technologies may be less standardised than in more established fields, requiring more detailed descriptions. Emerging technical fields rely more on recent science than older fields, and are based on (yet) less well known natural phenomena, which require more explanation than artefacts based on mechanisms recognised and accepted for long. This is notably the case for biotechnology, based on molecular biology, and of software, based on maths, algorithm and operational research. Furthermore, markets where technology is the most important competitive argument, where there is more licensing and cross-licensing, may encourage industry players to claim more and establish their rights with higher precision. Therefore, the sectoral specificity of a patent application might very well affect its voluminosity, and the surge in

---

[16] The 18 largest countries whose applicants have filed at least 10,000 EPO applications (about 0.5% of the total sample) over the entire period (1978-2004) were selected. The other countries are identified by the dummy variable "*APP_ROW*". Note that given their very high correlation with countries of priority, using the latter ones produces very similar results.

[17] According to Archontopoulos *et al.* (2006), the country of the applicant is a fairly good proxy for the language of filing (patentees tending to file their EPO applications in either their home language or English if they have no EPO language). But under Article 14 of EPC, it is possible for applicants to file their patents in non-EPO languages provided they supply a valid translation within 3 months. In some countries (especially Japan), applicants often tend (for up to 30 or 35% of their applications) to file their applications in their home language instead, making the applicant country-language correspondence less predictable, hence the need to control for non-EPO languages in our model.

[18] See van Zeebroeck *et al.* (2006) for an analysis of the impact of the choice of a classification on patent statistics.

patenting in new, science-based and extremely competitive technological areas might contribute to the increase in patent voluminosity at large.

This hypothesis is tested through 14 dummies representing the joint clusters used at the EPO to dispatch the applications to the proper examination departments. These dummy variables were computed from the IPC4 classes associated with the applications, through a matching table in use at the EPO. These clusters (JC) – listed in table 2 – represent broad technological areas, such as industrial chemistry (JC-01), telecommunications (JC-05) and human necessities (JC-14). Organic chemistry (JC-02), handling and processing (JC-11), and biotechnologies (JC-04) are the largest technological areas. Their shares in total EPO applications are 13.6, 12.5 and 12.1% respectively, while the computers and audio, video and media joint clusters are the smallest ones, each accounting for about 5% of all EPO applications.

H4: Patenting strategies

As the IP strategy of companies in certain industries has moved from being static (leveraging exclusion rights) to more active (trading rights, using them for licenses or other strategic purposes defined in Guellec et al. (2007)), their patent strategy has changed from a "single patent" view to a "portfolio management" view. It is well known that some large firms have a large propensity to patent their inventions. IBM for instance uses the fact it is the largest patentee in the USA as a marketing tool for several years and Microsoft has recruited an IP officer (from IBM) for developing its patenting strategy. In this context, what matters can then become the size and strength of the portfolio rather than the quality (scope, ability to stand in courts) of any single patent (see for instance Shapiro, 2000; Hall and Ziedonis, 2001; and Bessen, 2003).

The number of claims per patent could be the result, intended or not, of certain strategic choices of applicants, such as the ones detailed in section 2. Applicants claim more now than before in view of securing their position in new competitive conditions. One such strategy is the early patenting of yet unfocused inventions, which are narrowed down later in the process of examination. This could be identified through the use of divisionals: a divisional application is a sub-part of an initial application which does not satisfy the requirement of the EPC regarding the unity of the invention, but of which the applicant can secure the most important claims in a smaller application with a similar priority number and one or several further applications with the same priority date. One would expect that in the case of a divisionals strategy, the initial application, the "parent", would be bigger than average.

A second aspect may be the experience or lack of experience of the patentees with the patent system. In many cases, it will be easier to draft a longer application than a shorter one, resulting in applications with many unclear, often overlapping and redundant claims. This comes from under-staffed patent departments, or from the need to hire new, inexperienced staff that could not be properly trained. This is often not intended by the applicant, notably as it reduces the chances of having a quick grant and increases processing costs (due to more interactions with the examiners etc.), but it is a by-product of a strategy putting quantity over quality as a priority and may in some seldom cases be the effect a deliberate strategy to pollute a technological field or create a smoke screen around it.

These various aspects of patenting strategies are captured through a set of 5 variables: *PRIO* provides the number of priority patents claimed in the EPO application (about 1.2 on average); *HASDIV* identifies those applications that gave rise to subsequent divisional filings, which is the case of about 2% of the dataset; to the contrary, *ISDIV* isolates divisional filings themselves, representing about 2.4% of the filings. Finally, building on Kortum and Lerner (1999)'s approach, two variables are built up: *SIZE* gives the cumulative number of additional applications filed by the same applicant during the same year and the four consecutive previous years (420 filings on average) and *OCCAS* represents the inexperience of the applicant by marking those applications that are the only filing of their applicants in the current and the four previous years (about 23% of the filings). These two latter variables were computed using the official applicant codes from the main EPO database, i.e. without any cleaning of applicants' names.[19]

H0: The time trend

Next to all these potential factors, patents may become larger and larger simply as a consequence of global changes in economic environments, in market conditions, in courts behaviours and expectations or in writing and documenting habits, to mention only a few. This overall trend, extraneous to the above hypotheses, might be related to a general propensity toward increasingly lengthy and detailed technical descriptions in every field of human activities over time. Figure 1 precisely illustrates a clear trend component in the increasing voluminosity (the drop of the number of pages in the mid eighties is an artefact due to changes in the patent format at EPO). Therefore, it is important to capture in a way this potential effect of time and potentially non-accounted for the factors listed in the four main hypotheses. Table 2 further reflects the constant increase in the number of filings in recent years (the drop in 2004 is another artefact due here to the exclusion of the numerous international PCT filings not yet duly transferred to the EPO), as already observed by Kortum and Lerner (1999) for the United States.[20] To capture this trend, the model includes a set of 23 time dummies representing the year of filing at the EPO.

4. Empirical results

The model

The following model is used in order to test the four broad hypotheses on the determinants of voluminosity:

$$V_i = f(\beta_j, H_{ji}, T_i) + e_i , \qquad (1)$$

where V denotes the voluminosity indicators (number of claims or number of pages) for each patent *i* (*i*=1, …, 1,551,769). The endogenous variables are described in table 1. β_j are the vectors of parameters to be estimated. H_j are the vectors of explanatory variables, described in table 2 and summarized in table 3 corresponding to the four main hypotheses (*j*=1,…,4). *T* captures the trend factor (represented by time dummies) and *e* is the error term. Our dataset, created from different EPO databases,[21] comprises all Euro-direct and Euro-transferred PCT applications filed before the

---

[19] Following Trajtenberg (2004), this means that our data is subject to "type I errors" only, i.e. missing names that should go together, which leads to a likely underestimation of the variable *SIZE* and overestimation of the variable *OCCAS*. With other words, some patentees may have been erroneously declared occasional and recurrent patentees probably have larger portfolios in reality than accounted in our data.

[20] Kortum and Lerner (1999) suggest that the increase in patenting in the US has been driven by changes in the management of innovation of US firms which brought a real burst of innovation and an increased propensity to patent. Peeters and van Pottelsberghe (2006) provide evidence on the impact of these changes on the size of patent portfolios. These changes include an intensification of collaborative R&D, especially with universities; a focus on basic and applied research, an orientation towards product innovation, and the ability to reduce or overcome the traditional barriers to innovation.

[21] Including EPO (2006).

EPO between 1978 and 2004.[22] The estimates are run over the periods 1982-2004 for the number of claims and 1988-2002 for the number of pages (due to the formatting issues discussed here below and to the unavailability of page counts data for post 2002 applications).

Table 3 - List of exogenous variables

H1: national practices and US model diffusion hypothesis (23 variables):	
PCT	a dummy equal to 1 for PCT applications and 0 otherwise
USPR	a dummy equal to 1 for filings with a US priority applied by a non-US applicant
USAP	a dummy equal to 1 for filings with a non-US priority applied by a US applicant
APP_XX	19 country dummies (18 major countries of applicants + Rest of the world)
NO_EPL	a dummy equal to 1 for applications filed in another language than the three official EPO languages (DE, EN or FR) (mainly concerns Japanese applications filed in Japanese)
H2: the technological complexity hypothesis, It is composed of four variables	
INV	the number of inventors
IPC8	the number of 8-digit IPC classes listed
BPC	the number of backward patent citations
NPC	the number of non-patent backward citations
H3: the emerging sectors hypothesis, composed of 14 dummies (1 for each EPO Joint Cluster)	
H4: the strategic patenting hypothesis, composed of 5 variables	
PRIO	the number of priority applications listed in the patent filing
HASDIV	a dummy equal to 1 if the application led to one or more divisionals and 0 otherwise
ISDIV	a dummy equal to 1 for divisional applications and 0 otherwise
SIZE	the cumulative number of applications filed by the applicant in the same year and the 4 previous years on top of the current application
OCCAS	a dummy equal to 1 if the applicant had no other filing in the current and 4 previous years
H0: 23 time dummies (ranging from 1982 to 2004) to control for the effect of time	

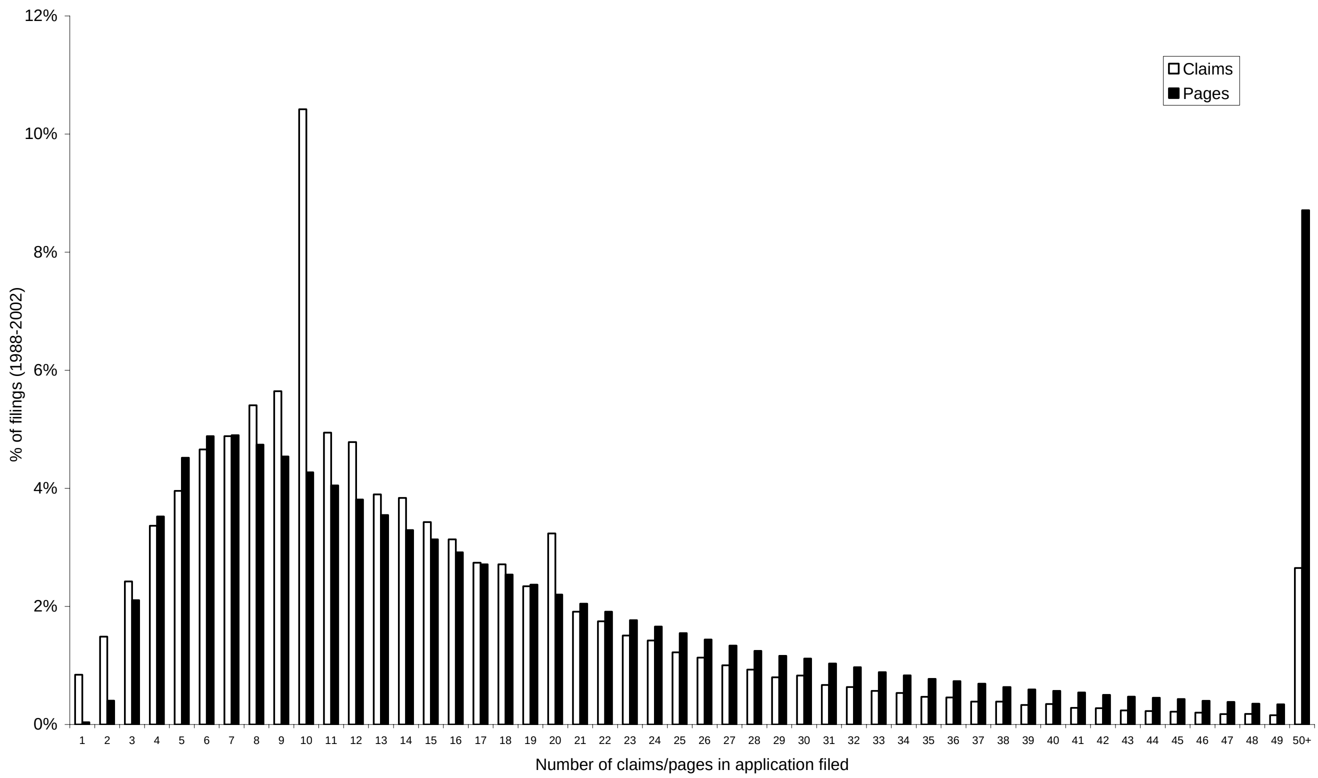
Our empirical methodology consists first in running subsequent estimates of equation (E1) with an OLS regression on the basic sample, with each independent variable and each individual hypothesis alone, and then with all the hypotheses simultaneously. The objective is to get first a broad idea of the explanatory power of each of the four hypotheses at large, then to estimate the the simultaneous impact of the determinants included in each of the four hypotheses, and finally to calculate the contributions of the various factors at the aggregate level. To perform the two latter steps, the count nature (i.e. discreteness and non-negativeness) and high skewness depicted by the distribution of the dependent variables (see figure 2), dictate the recourse to a count model with a negative binomial specification (see Hausmann, Hall and Griliches (1984) and Cameron and Trivedi (1986)).[23] Hence, we assume that the number of claims and pages is an exponential function of the variables listed above, so that the general form of the log-linear regression model specification would be:

[22] See Khan and Dernis (2005) for an interesting discussion on the impact of including or excluding international phase PCT applications on EPO statistics.
[23] The sample variance of the number of claims is about 146, and it is about 1292 for the number of pages, indicating substantial overdispersion in raw terms in both cases.

$$E\left[V_i \middle| H_i, T_i\right] = \lambda_i = \exp\left(\sum_{j=1}^{4} \beta_j H_{ji} + \beta_t T_i + e_i\right) \tag{2}$$

The error term e represents unobserved variables and measurement errors on the data and is assumed to be Gamma distributed with parameters $1/\alpha$ where α is the overdispersion parameter. The dispersion for the i-th observation is a function of the expected mean of the counts for this observation, i.e. $1+\alpha\lambda_i$. The model is estimated using maximum likelihood for the negative binomial distribution. The reported likelihood ratio test for overdispersion (α) rejects the null hypothesis of equidispersion, which confirms the preference for the negative binomial regression model over the pure Poisson one.

Figure 2 – Frequency distribution of claims and pages (1988-2002)



Source – own calculations based on EPO data

Figure 2 displays the frequency distribution of the endogenous variables (claims and pages). At first look, it reveals a high skewness of both distributions, with a very long upper tail, especially with page counts. Moreover, a strong institutional bias appears in the distribution of claim numbers in the form of an absolute mode of the distribution at 10 claims. This bias is due to EPO's excess-claim fees, making an additional fee due as from the eleventh claim on, which suggests some kind of price elasticity of the number of claims to claim-based fees. An analogous though weaker effect can furthermore be observed at 20 claims, which corresponds to the equivalent USPTO limit. Since the objective of this exercise is more to understand why applicants file claims in excess of the institutional threshold, the estimates are run on a transformed number of claims, shifted by ten units to the left according to equation (3), with M equal to 10.[24]

[24] Note that this transformation further increases the overdispersion of the distribution, reinforcing the choice of a negative binomial specification over a pure Poisson one.

$$V_i' = \max(V_i - M; 0) \qquad (3)$$

The number of pages has been perturbed by some external factors, as discussed by Archontopoulos *et al.* (2006). The most critical issue comes from the fact that EPO-Direct and PCT applications have different formats. Since the mid-eighties, when direct applications have been filed at the EPO, the description and claims sections of the applied documents have indeed been computerized and reformatted by the Office into a standard highly compact template, also known as '*type-set*' format. This reformatting of direct EPO applications, with the same font and layout for all EPO applications, makes the number of pages of these documents highly comparable with each other. To the contrary, PCT applications have not been reformatted by the EPO and the number of pages in the database corresponds then to the original (and heterogeneous) facsimile documents, as it can be observed in figure A1 in the appendix. Because of these formatting issues, the number of pages can hardly be analyzed for pre-1988 applications and must be controlled for the filing route as of 1988. The *PCT* variable should therefore capture this structural effect and is granted a higher explanatory power on the number of pages than it really deserves. Here also, in order to focus on pages in excess of the distribution mode, the model in (2) is nevertheless estimated with a transformed number of pages, according to (3) with *M* equals to 6.

Various robustness tests have been performed, largely confirming the results presented below: estimating the same model with OLS regressions, with priority countries instead of applicants countries, dropping variables, running the estimates (both negative binomial and OLS) on small random samples or different sub periods, performing separate regressions by country and by sector as well as for PCT *versus* Euro-Direct filings or for grants versus applications, and using priority countries instead of applicants' countries. In addition, the same model with a slightly reduced set of variables has been run on an extended dataset, including international PCT filings for which the EPO was designated as the International Search Agent (ISA) (1,931,631 filings in total), in which the international or regional status of PCT applications was also controlled for.[25] Finally, to evaluate the impact of the formatting issues evoked here above on the estimates of the number of pages, the same model has been used to explain the number of evenly-formatted granted patents in parallel with the number of pages in the corresponding initial applications. This latter test shows that this formatting issue does not distort the results presented here. More generally, all these tests are in line with and supportive of the following results.[26]

The explanatory power of the four hypotheses

The adjusted R-squared of the OLS estimations with the four broad hypotheses individually are presented in table 4. It clearly shows that each hypothesis and the model at large are stronger at explaining the number of pages than the number of claims.

At first sight it appears that the most relevant hypotheses are the "national practices and internationalization of the US model" and the "technological complexity" hypotheses (*H1* and *H2,* respectively) followed by the "emerging sectors" and "patenting strategies" hypotheses (*H3* and *H4,* respectively). When explaining the number of claims, the first hypothesis has a very high explanatory power (with an adjusted R-squared of about 11%). The technological complexity hypothesis explains about 5% of the variance and the emerging sectors, patenting strategies and

[25] i.e. without backward citations in the complexity hypothesis and without the *ISDIV* variable in the patenting strategies hypothesis for these variables were not available for international PCT applications.
[26] The results of all the robustness estimates are available upon request.

trend, about 3% each. For the number of pages, although the ranking is the same, hypotheses 2, 3 and 4 have more explanatory power than for the claims, with 11% for the former and 7% for the two latter. However, it is important to bear in mind that the '*PCT*' variable within *H1* captures the effect of the reformatting issue evoked here above and hence is granted much stronger a power than it probably deserves. Nevertheless, even when separating PCT applications (in heterogeneous facsimile formats) from Euro-Direct applications (in type-set format) within clustered regressions or when looking at the evenly type-set formatted grant publications, the explanatory power of *H1* on the number of pages remains very high (with an adjusted- R-squared of 0.09 in the latter regression).

Table 4 – Explanatory power of the 4 hypotheses
(Adjusted R^2 of the OLS models)

Hypothesis	Claims	Pages
H1: Internationalisation and US diffusion	0,11	0,12
H2: Technological complexity	0,05	0,11
H3: Emerging sectors	0,03	0,07
H4: Patenting strategies	0,03	0,07
H0: Trend	0,03	0,02
H1+H2	0,14	0,20
H1+H2+H3	0,15	0,22
H1+H2+H3+H4	0,17	0,25
H1+H2+H3+H4+H0	0,18	0,26

Estimates run on the same sample for claims and pages counts
Period: 1988-2002 - 1,092,164 observations

All this suggests the following preliminary observations:

- The model looks better at explaining the number of pages than the number of claims.
- The first hypothesis (H1) has the strongest explanatory power, suggesting at first sight that country-specific features (languages, patent cultures, drafting modes, parameters of the patent system) play the most important role on the size (claims and pages) of patent applications at EPO.
- The technological complexity comes clearly second in both cases, followed by the emerging sectors and patenting strategies at equal distance.
- Finally, the trend is weaker than the hypotheses, albeit slightly stronger with the number of claims than pages.

When all the hypotheses are included in the model, the estimates explain about 18% of the variance in the number of claims and 26% of the variance in the number of pages.

Maximum likelihood estimates for the negative binomial distribution are reported in table 5, which displays the estimated parameters for both claims and pages counts.

National practices and diffusion of the US model

The results provide strong support for the "*national practices and diffusion of the US model*" hypothesis. Indeed, with France as the reference, the US applicant dummy is associated with a very large and significant parameter, one of the most significant parameters of the model, suggesting that a patent filed at the EPO by a US applicant is composed of four additional claims and seven

Table 5 – Econometric estimates of equation E1 (Negative Binomial Regression)

	Claims 1982-2004				Pages 1988-2002			
	Coef.	ε (1)	**z**		**Coef.**	ε (1)	**z**	
H1: Internationalisation of application routes and diffusion of US modes (Reference for countries = France)								
PCT Filing	0,27	1,25	70,49	(**)	1,29	14,78	520,17	(**)
Non US Applicant with US priority	0,68	4,26	77,83	(**)	0,37	4,10	62,03	(**)
US Applicant with non US priority	-0,57	-1,98	-46,88	(**)	-0,35	-2,81	-44,94	(**)
AT Applicant	-0,07	-0,31	-4,47	(**)	-0,19	-1,59	-16,36	(**)
AU Applicant	0,58	3,53	30,57	(**)	0,38	4,34	29,91	(**)
BE Applicant	0,16	0,78	9,76	(**)	0,33	3,67	29,62	(**)
CA Applicant	0,38	2,09	24,69	(**)	0,38	4,36	36,54	(**)
CH Applicant	0,23	1,17	24,28	(**)	0,17	1,71	25,35	(**)
DE Applicant	-0,04	-0,18	-6,29	(**)	-0,09	-0,85	-21,19	(**)
DK Applicant	0,30	1,54	14,95	(**)	0,36	4,13	28,37	(**)
ES Applicant	-0,68	-2,23	-28,69	(**)	0,02	0,23	1,46	
FI Applicant	0,10	0,47	5,99	(**)	0,10	0,94	8,81	(**)
GB Applicant	0,30	1,55	35,78	(**)	0,16	1,60	27,80	(**)
IL Applicant	0,51	2,96	22,30	(**)	0,59	7,59	40,19	(**)
IT Applicant	-0,03	-0,14	-3,17	(**)	0,10	1,03	14,15	(**)
JP Applicant	0,05	0,23	7,53	(**)	0,99	13,20	214,73	(**)
KR Applicant	0,19	0,92	10,07	(**)	0,50	6,00	38,67	(**)
NL Applicant	-0,12	-0,50	-12,21	(**)	-0,07	-0,62	-10,38	(**)
SE Applicant	0,04	0,18	3,18	(**)	0,18	1,86	22,07	(**)
US Applicant	0,77	4,16	125,06	(**)	0,67	7,36	160,31	(**)
Applicant from the ROW	0,08	0,39	7,18	(**)	0,19	1,96	24,38	(**)
Filing in a non-EPO (DE/EN/FR) language					-0,68	-4,92	-147,60	(**)
H2: Technological complexity								
# Inventors	0,06	0,26	61,93	(**)	0,10	0,98	170,17	(**)
# IPC-7	0,07	0,33	58,41	(**)	0,09	0,85	115,62	(**)
# Backward Patent Citations	0,05	0,22	87,41	(**)	0,01	0,14	41,63	(**)
# Non Patent Citations	0,04	0,17	40,06	(**)	0,04	0,36	67,10	(**)
H3 Emerging sectors (Reference = JC-02 - Organic Chemistry)								
JC-01 - Industrial Chemistry	0,03	0,13	5,34	(**)	-0,35	-2,86	-99,28	(**)
JC-03 - Polymers	0,07	0,34	13,26	(**)	-0,01	-0,13	-3,66	(**)
JC-04 - Biotechnology	0,31	1,55	53,34	(**)	0,40	4,37	105,86	(**)
JC-05 - Telecommunications	0,29	1,46	37,23	(**)	0,07	0,66	13,74	(**)
JC-06 - Audio/Video/Media	0,35	1,85	45,90	(**)	0,21	2,13	42,21	(**)
JC-07 - Electronics	0,04	0,19	6,73	(**)	0,02	0,16	4,16	(**)
JC-08 - Electricity & Elec. Machines	0,00	0,02	0,85		-0,22	-1,87	-59,75	(**)
JC-09 - Computers	0,33	1,72	42,33	(**)	0,39	4,42	77,39	(**)
JC-10 - Measuring Optics	0,08	0,39	15,17	(**)	-0,11	-1,01	-30,39	(**)
JC-11 - Handling & Processing	-0,02	-0,09	-3,72	(**)	-0,23	-1,96	-63,76	(**)
JC-12 - Vehicles & Gen. Technology	-0,21	-0,87	-36,36	(**)	-0,34	-2,77	-86,20	(**)
JC-13 - Civil Engineering / Thermodynamics	-0,15	-0,65	-25,16	(**)	-0,30	-2,49	-72,22	(**)
JC-14 - Human Necessities	0,08	0,39	15,67	(**)	-0,17	-1,47	-45,82	(**)
H4: Patenting strategies								
# Priorities	0,25	1,14	115,72	(**)	0,21	1,94	146,30	(**)
Application has issued divisionals	0,62	3,78	58,28	(**)	0,38	4,35	61,14	(**)
Application is a divisional	-0,22	-0,89	-20,51	(**)	0,59	7,43	87,06	(**)
Cumulative Filings (Coef. x10e3)	-0,06	-0,26	-37,39	(**)	-0,01	-0,05	-4,97	(**)
Occasional (no filing in 5 prev, yrs)	0,10	0,48	27,01	(**)	0,02	0,18	7,22	(**)
Year of filing (Reference = 1988)								
1982	-0,23		-15,68	(**)				
1983	-0,18		-12,55	(**)				
1984	-0,14		-10,61	(**)				
1985	-0,12		-8,96	(**)				
1986	-0,10		-8,10	(**)				
1987	-0,06		-4,51	(**)				
1989	0,04		3,29	(**)	0,01		1,37	
1990	0,06		5,10	(**)	0,04		5,81	(**)
1991	0,07		5,70	(**)	0,08		11,75	(**)
1992	0,05		4,24	(**)	0,10		15,14	(**)
1993	0,10		8,42	(**)	0,12		18,44	(**)
1994	0,15		12,75	(**)	0,14		20,34	(**)
1995	0,17		14,88	(**)	0,13		19,42	(**)
1996	0,23		20,74	(**)	0,16		23,96	(**)
1997	0,30		27,83	(**)	0,17		25,70	(**)
1998	0,36		33,85	(**)	0,17		27,37	(**)
1999	0,45		42,65	(**)	0,20		31,33	(**)
2000	0,52		49,64	(**)	0,23		37,72	(**)
2001	0,60		57,46	(**)	0,25		40,88	(**)
2002	0,66		62,67	(**)	0,32		51,23	(**)
2003	0,72		68,60	(**)				
2004	0,76		63,74	(**)				
Constant	-0,08		-7,05	(**)	0,51		72,05	(**)
F-Stat / Log likelihood	-3,53E+06				-3,62E+06			
LN(alpha) [Std.Err.]	1,14 [0,002]				-0,01 [0,002]			
LR Test of alpha=0	10,00E+06 (**)				11,00E+06 (**)			
# Observations	1 454 552				1 092 164			

Significativity level: (*) p < 5% - (**) p < 1%

(1) Marginal elasticities computed for a hypothetic patent characterized by all explanatory variables equal to their average value at dy/dx. Takes into account discrete change of dummy variables from 0 to 1

additional pages than the average patent application at the EPO. To a lesser extent, other Anglo-Saxon countries also outfit the voluminosity of patent applications at the EPO: patents filed by British, Australian and Canadian applicants include on average one to 3 additional claims and one and a half to 4 additional pages. Some smaller countries such as Israel, Denmark, Korea and Switzerland show similar properties. On the contrary, continental European countries such as Germany, Spain and the Netherlands tend to have less claims and pages. Furthermore, a non-reported regression in which all exogenous variables were interacted with a time trend, allowing the observation of changing effects over time, shows that the value and significance of the country effects have strongly increased over the entire period of observation.

Even more support for the international diffusion hypothesis can be found in the estimated parameter of the *USPR* and *USAP* variables, indicating that patents filed with a US priority and then forwarded to the EPO by non US applicants are significantly larger both in terms of claims and pages (with four more claims and pages than the average). This tends to suggest that it might not be the culture of American applicants, but rather the American patent culture and system at large that induce a higher voluminosity.

Of particular interest in this respect is the case of Japan, since Japanese patents are not particularly composed of more claims but significantly much more pages. In other words, Japanese applicants seem to ask for fewer claims than many other countries' applicants, but include longer descriptions, possibly disclosing more of their inventions. What is more, the *NO_EPL* variable, identifying documents in a non-EPO language (i.e. the number of pages refers to the original PCT publication in another language than English, French or German instead of the EPO one), is associated with one of the most significant parameters of the model, inducing five pages less than the average. As this variable mainly captures Japanese PCT filings published in Japanese, this tends to suggest that the translation of Japanese patents into an official EPO language turns into much larger documents than their original counterparts. Beside the potential effect of language differences in characters and wording space, one possible interpretation could be that the assumed Japanese practice consisting in merging several national priorities into one single filing to the EPO or USPTO does result in files with more pages but not more claims than the average EPO filing.

On top of these national specificities, the PCT dummy is associated in both models with one of the largest and most significant coefficients (by far the most significant in the pages count regression). A patent filed under the PCT option contains indeed a bit more than one claim and almost 15 more pages than the average EPO application. This is highly supportive of the assumption that one dominant drafting style diffuses internationally through the PCT process (although once again the PCT variable is also capturing the format differences evoked here above for the number of pages).[27] This effect appears to be growing regarding the number of claims (i.e. its coefficient has increased over time) and constant for the number of pages.

Technological complexity

The four variables composing the technological complexity hypothesis all appear highly significant and positive determinants of patent voluminosity. In particular, four additional inventors induce one additional claim and four additional pages, suggesting that inventions originating from larger teams are broader and require longer descriptions. Similarly, the number of 8-digit IPC classes characterises potential architectural or complex inventions, which translates into applications with a higher voluminosity. An application linked to 3 more IPC classes is composed of one more claim

[27] As a reference, this variable also adds slightly more than one page to granted publications the format of which is uniform no matter the filing route.

and 2.5 more pages. Nevertheless, the influence of the number of inventors on patent voluminosity has strongly increased over time whereas the effect of the number of IPC classes has remained constant. Finally, the number of backward patent and non patent citations is also positive and significant in both models. Relying on more patented prior art induces slightly more claims (1 claim for 5 additional citations) and pages (1 page for 10 citations) whereas relying on the scientific literature adds one claim every 6 citations and 1 page every 3 citations.

Emerging sectors

The sectoral specificities of patent applications also have a strong impact on the voluminosity. With the organic chemistry cluster as the reference, the biotechnology area has the most significant and one of the largest positive effects on the number of claims and pages. On average, a patent application in the biotechnology cluster holds one and a half more claims and more than four additional pages. Of course, genetic sequences included in biotech filings almost surely play a role in this larger number of pages, but biotech applications also present more textual pages.[28] The audio, video and media cluster as well as the computers cluster (and to a lesser extent the telecommunications cluster) present similar effects on both the number of claims and pages. These sectoral effects show in addition a strong trend component when they are interacted with a time trend, suggesting that they have become stronger over time. For the remaining technological areas, some are associated with more claims but with less pages, such as industrial chemistry (with the strongest negative effect on pages), polymers, measuring optics and human necessities. Other sectors are associated with less claims and pages, namely vehicles and civil engineering.

Patenting strategies

In terms of patenting strategies, the one consisting in merging several priorities to file a single EP application leads to one of the strongest effects on both claims and pages. One additional priority leads indeed to one additional claim and almost two additional pages, probably witnessing that individual priorities may be copy-pasted into a new document. Of great interest are the two variables (*HASDIV* and *ISDIV*) relating to the filing of divisional applications. As logically expected, applications resulting into multiple divisional filings are much larger on average, with about 4 additional claims and pages. But surprisingly enough, those divisional applications themselves have only one fewer claim for even more pages, which suggests that applicants in such a case tend to drop claims but not the state of the art or the description of their invention from the initial filing.

Finally the experience (or lack of experience) of the applicants also influences the drafting of applications, but more significantly in terms of claims than pages, suggesting that it is more in the way they claim protection for their inventions that small and large applicants differ. In particular, larger applicants tend to file patents with fewer claims whereas occasional ones include slightly more claims into their applications.[29] This suggests that more experienced patentees have a capacity to draft their applications in a more focused way while accepting some rules of the disclosure game. To the contrary, applicants with less experience or with fast growing patent portfolios tend to claim for more exclusivity rights in a less synthetic way.

---

[28] As confirmed by additional non-reported regressions of the number of textual pages alone (i.e. illustrations excluded) as well as of the full number of pages with an additional control variable to account for the presence of drawings in the document.

[29] Note however, that including also the square of the *SIZE* variable into the regression highlights the non-linearity of the relationship since the square gets associated with a positive coefficient and the nominal variable a negative one.

Time trend

The trend effect appears very clearly in both estimates, especially for the claims where the time dummies for the early eighties have a negative and significant impact (1988 being the reference year). This negative impact decreases over time and becomes positive in the mid-nineties. A linear trend is highly visible, in terms of both the size of the parameters and their significance. This suggests that there remain some unaccounted for factors, extraneous to our hypotheses that influence the race toward larger and larger patents.

Contributions to the growth

The regressions above have identified the determinants of the size of each individual application. It is tempting to use these results in order to understand the dynamics of the average size of applications over the period 1980 to 2000. The coefficients reported in table 5 give a measure of the importance of the related factors and of their contribution to the number of claims and pages. Nevertheless, a finer measure of these contributions to the growth in voluminosity may be computed at the aggregate level based on the average number of claims or pages in year *t*:

$$\overline{V}_t = \varepsilon_j F(j,t) + \varepsilon_t + c + e_t \tag{4}$$

Where ε_j are the average elasticities computed for the average filing, *c* is an intercept and *F(j,t)* is the average of variable *j* in year *t* over the entire population, for instance the share of PCT applications in the total (i.e. the share of applications for which the PCT dummy takes the value 1) or the average number of inventors per application.

The growth of the number of claim or pages between year t_1 and year t_2 therefore writes as follows:

$$\overline{V}_{t2} - \overline{V}_{t1} = \varepsilon_j (F(j,t_2) - F(j,t_1)) + (\varepsilon_{t2} - \varepsilon_{t1}) + (e_{t2} - e_{t1}) \tag{5}$$

The contribution of each variable *j* is calculated as its elasticity ε_j multiplied by its average change between t_1 and t_2. The relative contribution of one factor, expressed in percentage, is its absolute contribution divided by the share of total change in the number of claims explained by the model, or:

$$RC_j = \frac{\varepsilon_j (F(j,t_2) - F(j,t_1))}{(\varepsilon_{t2} - \varepsilon_{t1}) + \sum_{j=1}^{J} \varepsilon_j (F(j,t_2) - F(j,t_1))} \tag{6}$$

Table 6 provides the contribution of each of the explanatory variables to the growth in the different voluminosity indicators according to equation (6). It suggests that few variables, even though they appeared as highly significant determinants of the voluminosity of individual filings, look strong contributors to the actual increase in the voluminosity of EPO filings. In terms of our hypotheses, the major contributor remains the diffusion of national specificities through the PCT route, which altogether contribute to about 15% of the growth in the number of claims and 60% in the number of pages. The PCT variable alone explains 10% of the claims' growth and 58% of the pages' growth (notably thanks to the formatting). But the US applicant dummy contributes 1% of the growth in both indicators and the non-US applicants with a US priority explain about 3% of the increase in claims and 1% of the increase in pages.

Table 6 - Contributions of endogenous variables to the growth in number of claims

	Claims 1982-2004				Pages 1988-2002			
Exogenous variables	**Nominal Growth**	ε (1)	**Contrib.**	**R.C.**	**Nominal Growth**	ε (1)	**Contrib.**	**R.C.**
H1: Internationalisation of application routes and diffusion of US modes (Reference for countries = France)								
PCT Filing	0,50	1,25	0,62	10%	0,44	14,78	6,44	58%
Non US Applicant with US priority	0,04	4,26	0,18	3%	0,04	4,10	0,15	1%
US Applicant with non US priority	0,00	-1,98	0,00	0%	0,00	-2,81	0,00	0%
AT Applicant	0,00	-0,31	0,00	0%	0,00	-1,59	0,00	0%
AU Applicant	0,00	3,53	0,00	0%	0,00	4,34	0,00	0%
BE Applicant	0,00	0,78	0,00	0%	0,00	3,67	0,00	0%
CA Applicant	0,00	2,09	0,01	0%	0,00	4,36	0,02	0%
CH Applicant	-0,01	1,17	-0,02	0%	0,00	1,71	0,00	0%
DE Applicant	-0,04	-0,18	0,01	0%	-0,02	-0,85	0,02	0%
DK Applicant	0,00	1,54	0,01	0%	0,00	4,13	0,01	0%
ES Applicant	0,01	-2,23	-0,01	0%	0,00	0,23	0,00	0%
FI Applicant	0,01	0,47	0,01	0%	0,01	0,94	0,01	0%
GB Applicant	-0,04	1,55	-0,07	-1%	-0,02	1,60	-0,04	0%
IL Applicant	0,01	2,96	0,02	0%	0,00	7,59	0,03	0%
IT Applicant	0,01	-0,14	0,00	0%	0,00	1,03	0,00	0%
JP Applicant	0,04	0,23	0,01	0%	-0,01	13,20	-0,11	-1%
KR Applicant	0,01	0,92	0,01	0%	0,01	6,00	0,09	1%
NL Applicant	0,00	-0,50	0,00	0%	0,00	-0,62	0,00	0%
SE Applicant	0,00	0,18	0,00	0%	0,00	1,86	0,01	0%
US Applicant	0,01	4,16	0,04	1%	0,01	7,36	0,10	1%
Applicant from the ROW	0,02	0,39	0,01	0%	0,01	1,96	0,02	0%
Filing in a non-EPO (DE/EN/FR) language					0,05	-4,92	-0,27	-2%
H2: Technological complexity								
# Inventors	0,63	0,26	0,17	3%	0,40	0,98	0,39	4%
# IPC-7	0,17	0,33	0,06	1%	0,16	0,85	0,14	1%
# Backward Patent Citations	0,56	0,22	0,12	2%	0,77	0,14	0,10	1%
# Non Patent Citations	0,49	0,17	0,08	1%	0,20	0,36	0,07	1%
H3 Emerging sectors (Reference=JC-02 - Organic Chemistry)								
JC-01 - Industrial Chemistry	-0,04	0,13	-0,01	0%	-0,03	-2,86	0,08	1%
JC-03 - Polymers	-0,03	0,34	-0,01	0%	-0,03	-0,13	0,00	0%
JC-04 - Biotechnology	0,05	1,55	0,08	1%	0,03	4,37	0,12	1%
JC-05 - Telecommunications	0,06	1,46	0,09	1%	0,06	0,66	0,04	0%
JC-06 - Audio/Video/Media	0,03	1,85	0,05	1%	0,02	2,13	0,05	0%
JC-07 - Electronics	0,00	0,19	0,00	0%	0,00	0,16	0,00	0%
JC-08 - Electricity & Elec. Machines	-0,01	0,02	0,00	0%	-0,01	-1,87	0,02	0%
JC-09 - Computers	0,04	1,72	0,07	1%	0,04	4,42	0,17	2%
JC-10 - Measuring Optics	0,00	0,39	0,00	0%	-0,01	-1,01	0,01	0%
JC-11 - Handling & Processing	-0,04	-0,09	0,00	0%	-0,04	-1,96	0,07	1%
JC-12 - Vehicles & Gen. Technology	0,00	-0,87	0,00	0%	0,00	-2,77	0,00	0%
JC-13 - Civil Engineering / Thermodynamics	-0,02	-0,65	0,01	0%	0,00	-2,49	0,01	0%
JC-14 - Human Necessities	0,01	0,39	0,00	0%	0,00	-1,47	0,00	0%
H4: Patenting strategies								
# Priorities	0,05	1,14	0,06	1%	-0,02	1,94	-0,03	0%
Application has issued divisionals	0,01	3,78	0,03	1%	0,00	4,35	-0,01	0%
Application is a divisional	0,01	-0,89	-0,01	0%	0,00	7,43	-0,02	0%
# Cumulative Filings (5 years) (Coef. *1000)	329,09	-0,26	-0,09	-1%	227,71	-0,05	-0,01	0%
Occasional (no filing in 4 prev, yrs)	-0,07	0,48	-0,04	-1%	-0,04	0,18	-0,01	0%
Year of filing (Reference=1988)								
Trend		4,92	4,92	77%		3,33	3,33	30%

(1) Marginal elasticities computed for a hypothetic patent characterized by all explanatory variables equal to their average value at dy/dx.Takes into account discrete change of dummy variables from 0 to 1

From the technological complexity hypothesis, contributing overall about 7% of the increasing voluminosity, the number of inventors remains clearly the most important factor, contributing to 3 to 4% of the increase in claims and pages. The other variables contribute between 1 and 2% each.

The emerging sectors, namely biotechnologies, telecommunications, medias and computers contribute together about 5% of the increase in claims and 3% of the increase in pages. The patenting strategy hypothesis contributes virtually not at all to the growth in voluminosity, except maybe for the number of priorities (*PRIO*) and the issuance of divisionals (*HASDIV*), which explain about 1% of the growth each. Although their coefficients have opposite signs, serial and occasional applicants are in fact both negative contributors to the increase in the number of claims.

5. Concluding remarks

The voluminosity of patent applications at the EPO has drastically increased over the past 20 years. The objective of this paper was to identify the sources of this phenomenon through an in-depth analysis of all patent filings at the EPO since 1982. The analysis relied on two indicators of voluminosity: the number of claims and the number of pages in each filing. It consisted in testing four broad hypotheses to explain what factors have influenced these indicators over the past two decades.

The four hypotheses (national practices and the diffusion of the US model, the technological complexity, emerging sectors and patenting strategies) all play some role in explaining the voluminosity of patent applications, but to different extents, and a larger one for the number of pages than for the number of claims. A significant trend effect also appears next to the hypotheses, especially with the claims.

That the patenting strategy hypothesis is associated with a high explanatory power suggests at first sight that there exist some patterns or habits in the way patents are drafted and applied which drive their scope and length. For instance, merging several national priorities into one EPO application does generate larger documents, presumably claiming for a broader protection, which suggests that in such a case, the original priority patents are simply summed up (with some potential redundancies) instead of being rewritten into a new more homogeneous and coherent application. Similarly, the increasing recourse to divisional applications reveals the more systematic early filing of unfocused inventions and also leads to larger documents. In addition, the experience or inexperience of the applicants in patenting shows at first glance a significant but relatively low impact on the voluminosity of their applications. Nevertheless, some patterns can be noticed in the drafting behaviour of the applicants, as new patentees tend to produce larger files, possibly due to a lack of experience in patenting, while more systematic patentees tend to claim for a more focused scope of protection in a more concise way.

Emerging sectors play a determinant role in the voluminosity of applications. Biotechnology patents, together with computer and audio, video and media patents are larger in terms of claims and pages than patents in other technological areas. To the contrary, industrial chemistry and polymers patents tend to have slightly more claims but less if not much less pages, while patents related to vehicles and civil engineering seem to be much smaller on average in terms of claims and pages. Such technological disparities may take their origin in the relative complexity of the related science, techniques and inventions, but could also be driven by industry-specific practices. As hypothesized earlier, this may also be related to the relative maturity or immaturity of the established vocabulary within the field, leading to more details and words being needed to fully cover the scope of a new invention.

The high explanatory power of technological complexity indicators and of the variables qualifying the nature of the invention may lead to the conclusion that some characteristics inherent to the organisation of research and to the invention process can also drive the size of the document that will be applied for a patent. The number of inventors and the number of IPC classes (at 8 digits) have a very high impact on the voluminosity of applications, indicating that inventions developed by larger teams and implying a wider range of technologies (i.e. architectural inventions) require more description and claims to be disclosed and protected. This effect could also be embodied in the way patent drafts are fed with information from the research teams, possibly implying every inventor to bring some content to the document. Similarly, a more intensive recourse to the state of the art (i.e., to the patent and scientific literature) generally induces larger applications. Since these indicators can be seen as proxies for the complexity of the inventions, this might also indicate that more complex inventions require more claims and more pages to be patented.

The fact that PCT filings tend to be so much larger in terms of both indicators can have various explanations and might reflect other factors. As the main benefit of the PCT option is to delay the costs to be incurred in the procedure while the invention gets more mature and the market clearer, it seems reasonable to consider that many PCT applications may be drafted somewhat in a hurry, without a precise view yet on the critical element to claim protection for, especially given the absence of claim-based fees at WIPO. Hence, since it is more difficult to reduce than to enlarge the claimed content of an application once the procedure is started, applicants in such a situation would draft applications in a broader, fuzzier and possibly longer way. The common recourse to fall-back positions or office-specific versions of a single claim can also contribute to explain this observed phenomenon.

Another explanation might be that PCT applications generally designate the USPTO and therefore are drafted according to the US system, for it represents the largest market and its granting process is usually shorter. In this respect, it is likely that the oversized characteristic of PCT applications is also reflecting some sort of geographical diffusion. In other words, PCT filings would follow a “draft once, file everywhere” principle, according to which patents are drafted with a US template then applied to all other patent offices. Having expended considerable resources drafting a patent application for one large market, there is a tendency to crib from it when making applications elsewhere. This idea is further strengthened by the geographical specificities observed from the analysis.

National differences in patenting behaviours or habits explain the heterogeneity of patent voluminosity to a very important extent. The US syndrome, demonstrated by much larger applications from US applicants (and from Anglo-Saxon countries in general) as compared to other (especially continental European) countries, is further confirmed by the larger size of patents applied by non-US applicants with a US priority instead of a national one (and oppositely the smaller size of patents applied by US applicants with a non-US priority). This suggests that the observed phenomenon most probably finds its roots in the nature of the American IP system at large, implying patentees applying to the US to describe and formulate their claims with much more breadth than what the continental European system would require.

The difference in size between North American and Continental European commercial contracts, auditor’s comments and other legal documents – anecdotically reported by some lawyers we interviewed – further supports the idea that the Common Law system induces larger documents than the Civil Law one does. This might be related to the traditionally larger role of precedents in US Courts’ decisions as compared to continental European Courts, to the civil law tradition precluding its judges from establishing broad principles of law in the absence of legislation – a possibility that is still at least theoretically open in the common law system –, and to litigation being much more common in the US than in Europe. Applicants to the US patent system may therefore have to be more exhaustive and detailed in their drafting so as to maximize the chances for their patents to stand in Court.

In addition, the US doctrine of equivalent, tempered by the process history estoppel, makes yet another very strong argument in favour of larger and more detailed applications, for what has been given up during the granting process can no longer be claimed back in Court. In addition, the mandatory best mode to be detailed in USPTO applications but not in EPO or JPO ones might be yet another contributor to this effect. It is indeed a frequent practice that American patentees detail several utilization modes to hide the "best" one, possibly leading to longer descriptions as well. And finally, the differences in fee regimes between USPTO and the EPO may constitute yet another factor behind this phenomenon. While the former applies extra fees for claims in excess of 20, the latter penalizes claims in excess of 10, hence the fee in the country of priority (and supposedly for which it is initially drafted) may dictate the length of the application when it is filed at the EPO, which is further supported by the frequency distribution of the number of claims, displaying a local mode at 20.

Taking into account this "American" or "Anglo-Saxon" effect embodied in the nature of PCT filings, the US diffusion hypothesis remains the most appealing one, leading to larger and more detailed documents, which does influence the size of patent applications at EPO, and does therefore contribute to the surge in the workload of EPO examiners.

Overall, the increasing number of claims and pages seems to result from the globalisation of technology markets with the diffusion of the US patent model, along with the increasing complexity of technologies, the technical and competitive processes in emerging sectors, a tendency to earlier patenting of unfocused inventions, and probably some additional factors not accounted for in the model. Indeed, there still remains a significant unexplained trend in the increase of the voluminosity of patent applications. This may be related to a general propensity to more complete, detailed, complex and hence voluminous literature in every field of activity, such as the user manuals of electronic devices, the documentation of mass-market consumer goods, official or technical reports, or even laws. Patents may be just another field of expression for this generalized volubility, encouraged by the decreasing costs of drafting and disseminating written information. Finally, this trend might also reflect strategic factors, related to intensified competition on markets (and in courts), which are not captured by our model.

References

Archontopoulos, A., D. Guellec, N. Stevnsborg, B. van Pottelsberghe de la Potterie and N. van Zeebroeck, "When small is beautiful: measuring the evolution and consequences of the voluminosity of patent applications at the EPO," *ULB, SBS-CEB and MIMEO Working Papers* (2006).

Bessen, J., "Patent Thickets: Strategic Patenting of Complex Technologies," *SSRN Working Paper* (2003). Available at SSRN: http://ssrn.com/abstract=327760.

Cameron, A. and P. Trivedi, "Econometric models based on count data: Comparisons and applications of some estimators and tests," in *Journal of Applied Econometrics*, Vol. 1(1), 29-53.

Dernis, H., D. Guellec and B van Pottelsberghe de la Potterie, "Using patent counts for cross-country comparisons of technology output," *OECD STI Review*, 27 (2001), 129-146.

_______ , "EPO Worldwide Patent Statistical Database," *European Patent Office*, April 2006.

Guellec, D. and B. van Pottelsberghe de la Potterie, “Applications, grants and the value of patents,” *Economic Letters*, Vol. 69 (1) (2000), 109-114.

Guellec, D. and B. van Pottelsberghe de la Potterie, “The internationalisation of technology analysed with patent data,” *Research Policy*, Vol. 30(8) (2001), 1256-1266.

Guellec, D. and B. van Pottelsberghe de la Potterie, “The value of patents and patenting strategies: countries and technology areas patterns,” *Economics of Innovation and New Technologies*, 11(2) (2002), 133-148.

Guellec, D., B. van Pottelsberghe de la Potterie and N. van Zeebroeck, “Patent as a market instrument,” in Guellec, D. and B. van Pottelsberghe de la Potterie, “The Economics of the European Patent System: Innovation, Competition, and Growth,” Oxford, *Oxford University Press*, forthcoming (2007), Chapter 4.

Hall, B., S. Graham, D. Harhoff and D. Mowery, “Prospects for Improving U.S. Patent Quality via Post-grant Opposition,” EconWPA, *Law and Economics*, No. 0401002 (2004).

Hall, B. and R. Ziedonis, “The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry 1979–1995,” *RAND Journal of Economics*, Vol. 32 (1) (2001), 101-128.

Hausman, J., B. Hall and Z. Griliches, “Econometric models for count data with an application to the patents – R&D relationship,” *Econometrica*, Vol. 52 (1984), 909-938.

_______, “Human Development Report 2001,” New York, *Oxford University Press, United Nations Development Programme* (2001).

Jaffe, A. and J. Lerner, “Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It,” Princeton, *Princeton University Press* (2004).

Jensen, P., A. Palangkaraya and E. Webster, “Patent Application Outcomes across the Trilateral Patent Offices,” *IP Research Institute of Australia Working Paper (2005)*, 12-26.

Khan, M. and H. Dernis, “Impact of Patent Cooperation Treaty data on EPO patent statistics and improving the timeliness of EPO indicators,” *OECD STI Working Paper* (2005) No. 2.

Kortum, S. and J. Lerner, “What is behind the recent surge in patenting,” *Research Policy*, Vol. 28 (1999), 1-22.

Lanjouw, J. and M. Shankerman, “The quality of ideas: measuring innovation with multiple indicators,” *NBER Working Paper* (1999).

Peeters, C. and van Pottelsberghe de la Potterie, B., “Innovation Strategy and the Patenting Behavior of Firms,” *Journal of Evolutionary Economics*, Vol. 16 (1) – 2 (2006), 109-135.

Shapiro, C., “Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting,” *NBER Working Paper* (2000).

Stevnsborg N. and B. van Pottelsberghe de la Potterie, “Patenting procedures and filing strategies,” in Guellec, D. and B. van Pottelsberghe de la Potterie, “The Economics of the European Patent System: Innovation, Competition, and Growth,” Oxford, *Oxford University Press*, forthcoming (2007), Chapter 7.

Tong, X. and J. Frame, “Measuring National Technological Performance with Patent Claims Data,” *Research Policy* Vol. 23 (1994), 133-141.

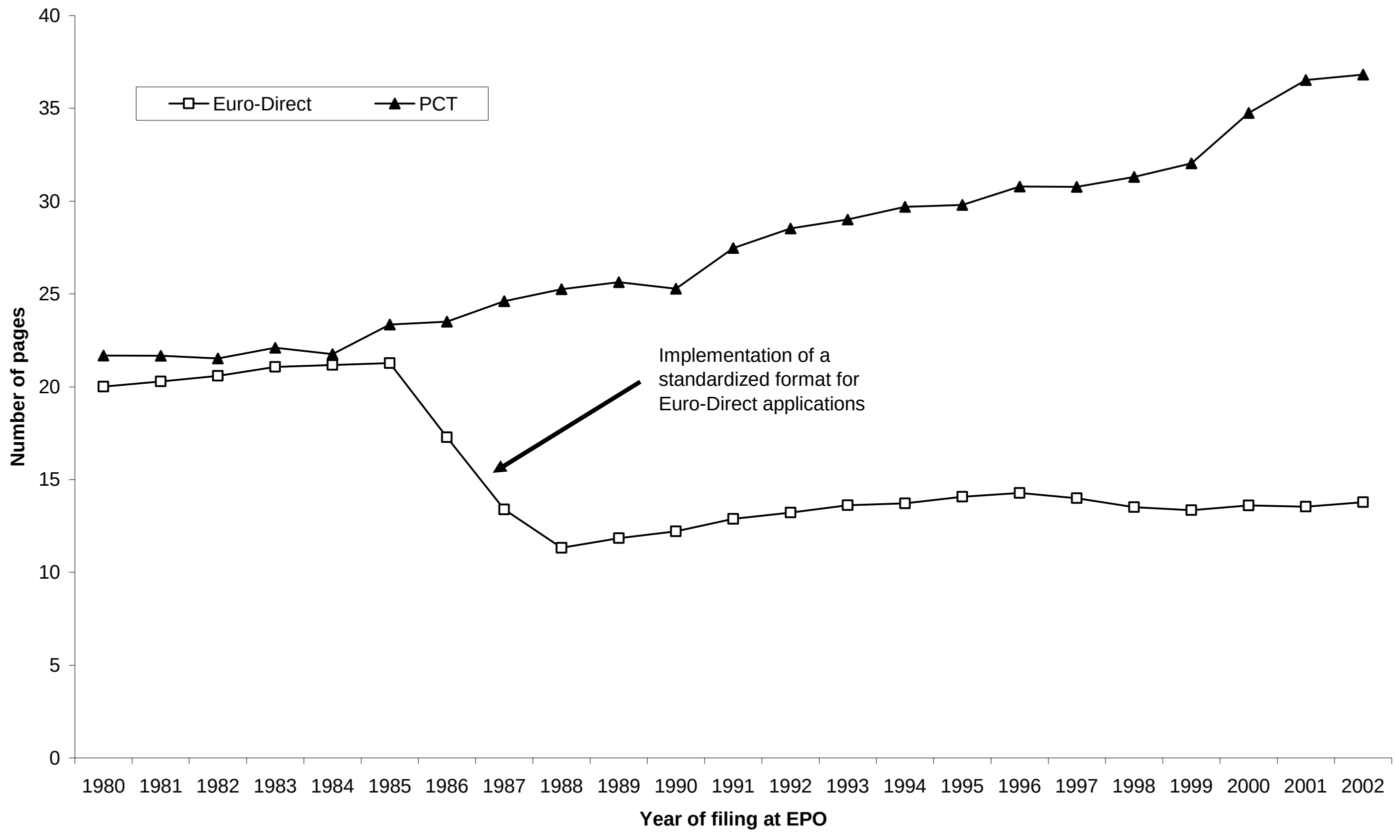
Trajtenberg, M., “The names game: tracing the mobility of inventors using patent data,” *Working paper presented at the Social Science & Technology Seminar*, Stanford Institute for Economic Policy Research (SIEPR), Stanford University, Stanford CA, November 1994.

van Pottelsberghe de la Potterie, B. and D. François, “The Cost Factor in Patent Systems,” *ULB, SBS, CEB Working Papers* (2006).

van Zeebroeck, N., B. van Pottelsberghe de la Potterie and W. Han, “Issues in measuring the degree of technological specialisation with patent data,” *Scientometrics* Vol. 66 (3) (2006), 481-492.

Appendixes

Figure A1 – Average number of pages in Euro-Direct and PCT applications at EPO (1980-2002)



Source – own calculations based on EPO data