ISSN 2315-0815

eurostat
Manuals and guidelines

Patent statistics at Eurostat: Mapping the contribution of SMEs in EU patenting

2014 edition

eurostat

eurostat
Manuals and guidelines

Patent statistics at Eurostat: Mapping the contribution of SMEs in EU patenting

2014 edition

eurostat

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*)**:**

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the Internet (http://europa.eu).

Luxembourg: Publications Office of the European Union, 2014

ISBN 978-92-79-39164-4
ISSN 2315-0815
doi:10.2785/55102
Cat. No: KS-GQ-14-009-EN-N

Theme: Science and Technology
Collection: Manuals and guidelines

Foreword

Dear readers,

I am pleased to introduce this edition of the Eurostat's Manuals and Guidelines series dedicated to the production and dissemination of enhanced statistics that measure SMEs' involvement in inventive activity in the European Union through patenting.

Investment in science, technology and innovation is one of the EU's central policy areas. It is an important driver for the Europe 2020 growth strategy and is essential to economic growth and the development of the knowledge-based economy. The Europe 2020 strategy sets out a vision of Europe's social market economy for the 21st century. It aims to turn the EU into a smart, sustainable and inclusive economy delivering high levels of employment, productivity and social cohesion.

Knowledge creation and innovation dynamics stem from the activities of a variety of actors, including firms, universities, entrepreneurs, and public and private research institutes. Patents are generally used to protect R&D results, but they are also valuable as a source of technical information without which ideas may need to be re-invented and re-developed.

Eurostat collects data in the areas of science, technology and innovation that are used both by policymakers and scientists. Patent statistics are recognised as a highly valuable data source for assessing innovative performance and for monitoring, evaluating and even forecasting firms' technological activities, regardless of their size.

Given SMEs' contribution in developing technology and high R&D productivity, it is very useful to establish the extent to which they are involved in innovative activities across countries. This publication assesses the feasibility of identifying SMEs' contribution to technological development by measuring their share of total patent activity. The challenge for the project was linking corporate patent applicants to business registers and then classifying patent applications according to firm size.

In terms of methodology, the project's contribution is that it demonstrates the feasibility of deriving SME patent indicators from SME shares in patent portfolios. It also points to potential future improvements by showing that the automated matching of patent applicants with information in business registers needs to be complemented with additional procedures to obtain accurate estimates and reliable statistics.

By showing how SMEs' involvement in EU technological activities can be monitored on the basis of relevant patent statistics, the methodology described in this publication marks a major step forward.

Maria-Helena FIGUEIRA,

Director of Global business statistics

Eurostat

Chief editor

Bernard Félix — Bernard.Felix@ec.europa.eu
Eurostat, Unit G4 — Innovation and information society

Editors

This report was prepared by Jan-Bart Vervenne ([2]), Julie Callaert ([1])([2]) and Bart Van Looy ([1])([2]) in collaboration with Sogeti Luxembourg S.A. (Gaëtan Châteaugiron and Frédéric Stibling).

Acknowledgements

The project was financed by the European Commission's Directorate-General for Research and Innovation.

The report benefited from expert input from the EUROSTAT Eurogroup Register on multinational business group membership of patenting companies and Machteld Hoskens ([1]) on the design of the adopted sample strategy (see Extrapolation section).

Further acknowledgements go to Caro Vereyen ([2]) for her extensive contribution to the online firm-size screening process and Xiaoyan Song ([1]) for the technical support.

For more information please consult

Eurostat
Statistical Office of the European Union
Bech Building
Rue Alphonse Weicker 5
L-2721 Luxembourg
Internet: http://ec.europa.eu/eurostat
E-mail : estat-user-support@ec.europa.eu

([1]) ECOOM, KU Leuven, Waaistraat 6 — box 3536, 3000 Leuven, Belgium.
([2]) INCENTIM, KU Leuven, Naamsestraat 69 — box 3535, 3000 Leuven, Belgium.

Contents

1. Introduction ([1])

Objective

Recent figures on the contribution of SMEs to European economic activity (Eurostat, 2011) reveal that an overwhelming majority (99.8 %) of the firms active in the European Union qualify as SMEs (figures for 2008) ([2]). Two in three jobs (66.7 %) stem from SME activity, and SMEs account for 58.6 % of value added.

SMEs' contribution to innovative activities has been the subject of extensive research since the early 20th century. Schumpeter (1911) was one of the first authors to highlight the importance of small enterprises in the innovation process. He identified entrepreneurs as key figures in the dynamics of 'creative destruction', since they turn inventions into innovations by creating enterprises to monetise marketable applications. Schumpeter (1942) assumed that, as the innovation process increasingly became routine, the role of the entrepreneur in the innovation ecosystem would become less important than that of monopolistic large firms.

Previous research relating innovation to firm size has revealed a number of robust empirical patterns. Schumpeter's (1942) expectations have been confirmed: R&D activities (measured by R&D expenditures or personnel) increase monotonically with the firm size of R&D actors. However Proportionally, R&D expenditures remain fairly constant regardless of firm sizeand SMEs display higher R&D productivity levels: they produce more patents than large firms per dollar spent on R&D (Cohen, 2010) ([3]).

Scholars have tried to reconcile the seemingly opposing views advanced by Schumpeter. Among others, Baumol (2002; 2004) nuanced the exclusive aspect of Schumpeter's (1911; 1942) view by emphasising the complementary roles of large incumbents and small entrepreneurial firms in the process of innovation in free market economies. Large oligopolistic firms engage with other incumbents in an R&D expenditure 'arms race', accumulating process innovations and incremental improvements to existing products in mature phases of the technology (and business) life-cycle. Independent innovators operating small business enterprises, on the other hand, account for many of the most revolutionary innovations in the past two centuries, innovations that have set in motion technological paradigm shifts (Baumol, 2004).

Baumol (2002; 2004) and Cohen (2010) provide a rationale that helps explain why large companies secure such a large share of incremental innovation and process innovation. Incumbents have a greater incentive to invest in incremental projects that exploit their existing R&D capabilities. Incremental innovations can magnify existing competitive advantage and strengthen the incumbent's market position. Incumbents' risk aversion leaves enterprising entrepreneurs plenty of scope to develop among others the ideas the former would deem too risky.

Scherer (1991), Rothwell (1989) and Audretsch (1995), *inter alia*, provide insights into the mechanisms by which SMEs introduce new products and services. Small firms have several advantages over large corporations that may help to explain their prevalence in shaping breakthroughs. According to Scherer (1991), the level of bureaucracy in most large firms is not conducive to risky R&D activities. In addition, 'disruptive' inventions can destroy cash flows, leading to further inertia at the level of (corporate) decision-making. In SMEs, by contrast, decisions can be made quickly and are largely unaffected by concerns relating to existing products and markets (OECD, 2000).

These theoretical conjectures and the empirical observation that SMEs can exploit R&D opportunities more efficiently and contribute relatively more to 'radical' or 'breakthrough' innovations underline the importance of assessing and monitoring the proportion of innovative activity for which they are responsible. In particular, a Europe-wide mapping of corporate patenting, broken down by firm size, can be used to evaluate and assess SMEs' contribution, and thus inform policy choices at EU and/or Member State level.

([1]) This report was prepared under a Eurostat contract.

([2]) Excluding the financial industry.

([3]) References to 'small firms', 'small enterprises' or 'small companies', as opposed to their large counterparts, implicitly cover all types of SME: medium-sized, small and micro.

Mapping the innovative contribution of SMEs by means of patents

The mapping of corporate patenting by firm size requires a classification of patent applicants, for which only name and address information is available in the patent databases. In the past decade, various techniques have been developed for analysing large quantities of patent data. Several patent data enhancements are relevant in this respect. Sector allocation methodologies (e.g. Van Looy *et al.*, 2008; Du Plessis *et al.*, 2011) help to identify firm applicants (not universities, hospitals, private and public non-profit organisations, governmental agencies and individuals). Name cleaning and harmonisation algorithms enable researchers to cope with different applicant names appearing in patent documents within and across patent systems.

To determine the role played by SMEs, one also needs data on firms' size and (in)dependency status. Previous studies have matched patent data to financial databases so as to be able to extract firm-size indicators from annual accounts (Hall *et al.*, 2001; 2005; Thoma *et al.*, 2010; Macartney, 2007). Some of these studies use the results to gauge patent activity by large enterprises on the one hand and SMEs on the other (Perrin & Speck, 2004; Iversen *et al.*, 2009; Helmers & Rogers, 2009; Frietsch *et al.*, 2012; Squicciarini & Dernis, 2012; CHI Research, 2003; Jensen & Webster, 2006; Keupp *et al.*, 2009). However, such studies tend not to distinguish between small subsidiaries of multinational enterprises and independent SMEs and/or they discard applicants for which available information is insufficient for determining size.

The research presented in this paper adds to this literature. In Section 2, we outline and apply a methodology for assessing SMEs' involvement in (patented) technology development in the EU. Our contribution fills a gap in the previous literature by complementing an automated methodology with additional, stratified search efforts for missing information in order to produce a comprehensive picture, and by distinguishing between small subsidiaries of multinational enterprises and independent SMEs. Section 3 sets out further analysis of the classified patent portfolios, including a less direct approach to evaluating SMEs' contribution to innovative activity based on relative technological advantages.

2. Methodology

The methodology used to derive reliable estimates of SMEs' contribution to corporate patenting in the EU consists of a number of steps, as follows:

i. corporate (patent) applicants are matched to financial directories;

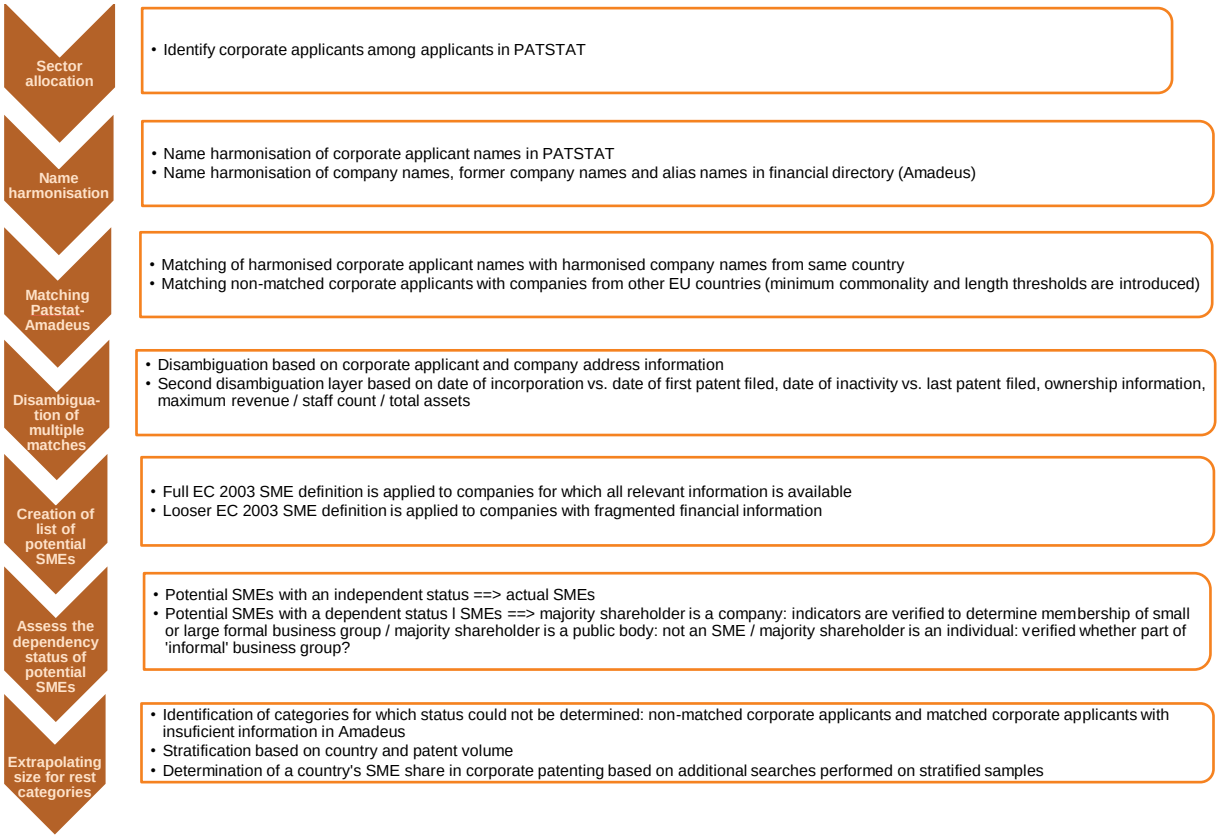
ii. a disambiguation procedure is applied to identify multiple companies that are matched to the same corporate applicants;

iii. relevant financial indicators and information on (in)dependency status enable us to map many applicants according to firm size; however, a non-negligible portion remains unidentified, due to information missing from the financial database, so;

iv. stratified samples of the corporate applicants that remain unmatched are investigated to assess firm size.

Figure 1 provides an overview of these steps.

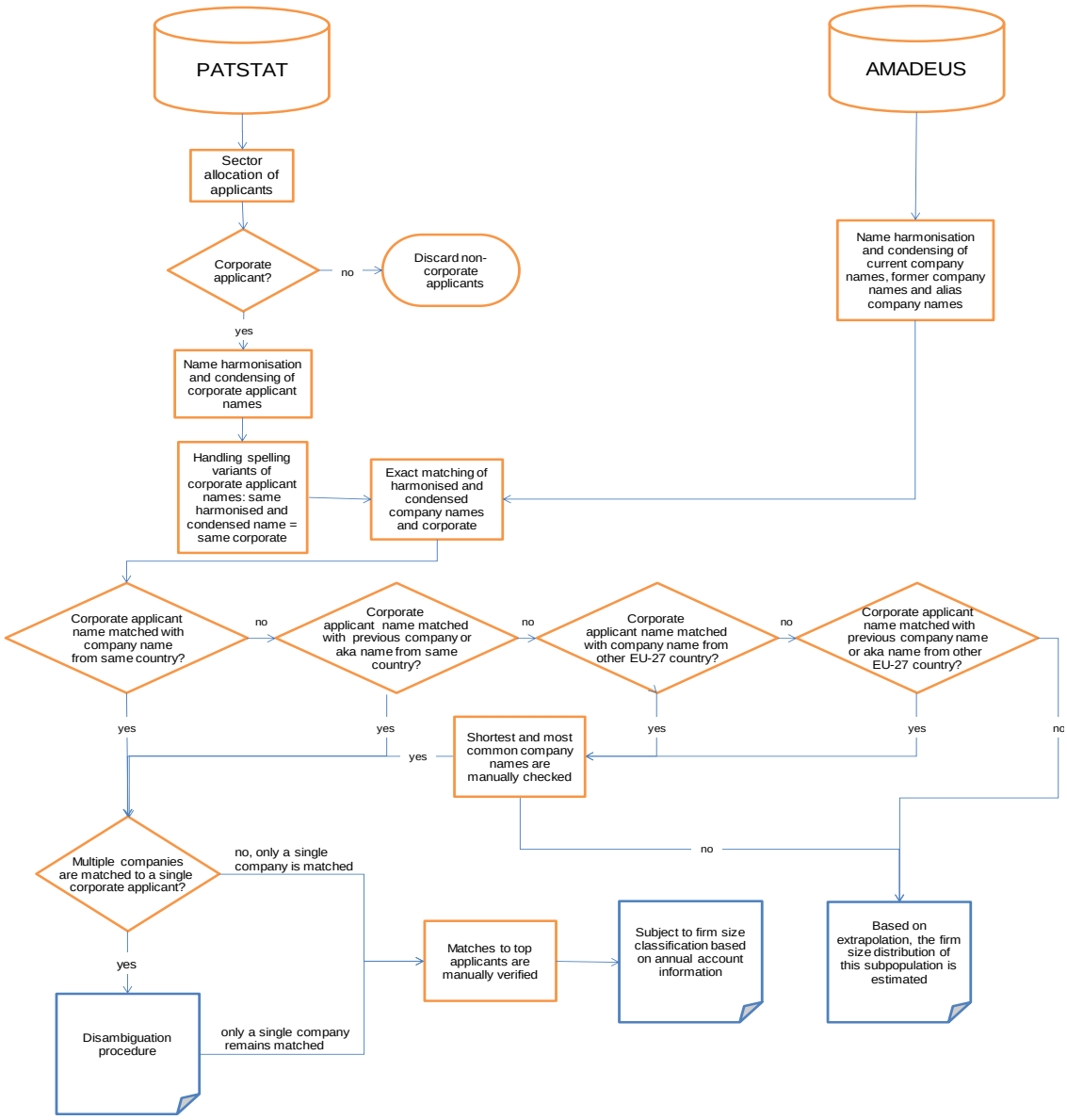
Figure 1: Flow chart of procedure to measure SMEs' contribution to corporate patenting in the EU



Matching

Patent databases contain only applicant names and address information, to the exclusion of additional information that would allow direct assessment of firm size and/or swift linkage with financial databases (such as a single company identifier). To classify corporate patent applicants in terms of size, therefore, we have to match applicant names with company names in financial directories from which the additional information can be extracted. We use patent data from the EPO Worldwide Patent Statistical Database or PATSTAT (autumn 2011 edition). Several databases contain financial data from annual account filings with national business registries. We use relevant firm-level information from across the EU from Bureau Van Dijk's (BvD) Amadeus database (2012 edition). Hence, our approach involves seeking correspondence between applicant names in PATSTAT and company names in Amadeus.

Figure 2: Process flowchart for matching procedure



In its raw form, the PATSTAT database provides unprocessed (non-harmonised) applicant names, as well as country and address information. Various procedures are therefore required to clean and enrich the raw patent data.

We applied a sector allocation algorithm to all applicant names in PATSTAT to limit the number of target applicants to be matched to financial directory records and reduce the odds of associating non-corporate applicants with companies. The algorithm uses a keyword logic to filter out non-corporate applicants (for more detail on the sector allocation methodology, see Van Looy *et al.*, 2011; du Plessis *et al.*, 2011) ([4]).

Next, using an automated matching approach, EU corporate applicant names in PATSTAT (filing for patents from 1999 onwards) are matched to names in Amadeus of firms established in the EU. In this

([4]) Other sectors seeking patent protection include individuals, government and non-profit bodies, and universities. Corporates accounted for 66 % of the patents filed in the countries in the reference period; individuals 29 %; governments and non-profit bodies 3 %; and universities 2 %.

study, we focus on corporate applicants filing for patent protection at EPO or USPTO, or relying on the PCT procedure. To limit the number of potential false negatives due solely to the presence of name variants in the databases, we applied the Magerman *et al.* (2006; updated in 2009) name harmonisation procedure to the lists of all company names in Amadeus and of corporate applicants in PATSTAT before the actual matching took place.

Discrepancies in company names relate to punctuation, legal form, spelling, characters and umlauts. Name harmonisation procedures are introduced to facilitate analysis at applicant level and ensure that patents filed by the same applicant are not classified as originating from a number of companies. We aggregated applicant counts at the level of the harmonised name, assuming that one harmonised applicant name in one country represented one business entity.

The actual matching consists of two rounds:

i. corporate applicants are matched exclusively to companies from the same country. Harmonised corporate applicant names are compared with harmonised current company names. For corporate applicants that remain unmatched, we compare harmonised former company names with company aliases ('also known as'); and

ii. unmatched corporate applicants are paired with companies from other Member States, on the assumption that subsidiaries may be established under names resembling the name of the parent company. As in the first round, names are compared with original company names, former company names and aliases, in that order.

The first round (country-by-country) minimises the number of multiple matches, while the second (all Member States) maximises the proportion of corporate applicants associated with a BvD company. Table 1 shows the success rates of both rounds for the entire time frame (application years 1999-2011) in terms of patent applicants and applications.

Table 1: Applicants and applications matched to at least one company in the financial directory

Country	Corporate applicants			Corporate applications		
	Total	Matched	%	Total	Matched	%
EU-27 ([1])	104 166	64 496	61.9	1 316 568	1 094 349	83.1
BE	2 218	1 542	69.5	26 129	23 220	88.9
BG	107	45	42.1	190	73	38.4
CZ	500	336	67.2	1 450	967	66.7
DK	3 593	2 101	58.5	29 487	24 468	83.0
DE	30 130	16 320	54.2	537 847	453 746	84.4
EE	112	65	58.0	226	136	60.2
IE	1 235	912	73.8	8 767	6 575	75.0
EL	209	59	28.2	676	196	29.0
ES	4 234	2 395	56.6	17 019	11 494	67.5
FR	10 763	5 587	51.9	179 457	144 112	80.3
IT	13 104	8 974	68.5	77 186	60 358	78.2
CY	245	62	25.3	932	323	34.7
LV	74	18	24.3	288	37	12.8
LT	16	8	50.0	27	13	48.1
LU	649	259	39.9	5 399	3 107	57.5
HU	513	181	35.3	1 689	636	37.7
MT	82	53	64.6	426	363	85.2
NL	6 891	4 720	68.5	132 865	121 315	91.3
AT	3 042	1 632	53.6	25 293	18 588	73.5
PL	401	238	59.4	1 179	796	67.5
PT	382	192	50.3	1 065	738	69.3
RO	57	17	29.8	95	34	35.8
SI	265	135	50.9	1 438	678	47.1
SK	124	76	61.3	305	225	73.8
FI	2 683	1 724	64.3	51 052	44 874	87.9
SE	6 226	3 452	55.4	84 844	53 081	62.6
UK	16 311	13 393	82.1	131 237	124 196	94.6

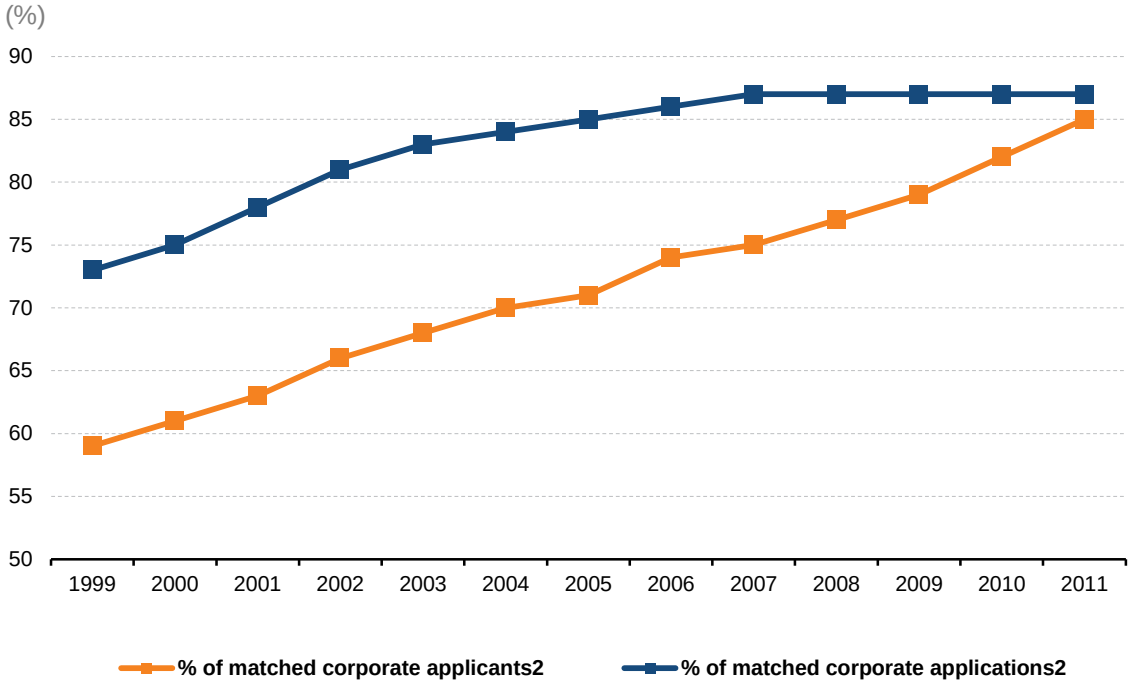
([1]) This study was carried out between 2011 and 2013. At that time study, the EU comprised 27 Member States. Therefore Croatia is not covered in this publication.

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

National matching rates (aggregated across patent offices) range between 24.3 % (Latvia) and 82.1 % (United Kingdom). Overall, 61.9 % of the harmonised corporate applicant names are matched to BvD companies (57.9 % in the same country and 4.0 % in other Member States). These matched corporate applicants account for 83.1 % of patent applications filed by corporate applicants. 77.9 % can be assigned to corporate applicants matched to companies from the same country and 5.2 % to those matched to companies from other Member States. A comparison between applicant and application figures shows that, on average, unmatched corporate applicants patent less than matched ones (the remaining 16.9 % of the corporate patent volume).

Figure 3 shows matching rates by application year. Overall, a trend is evident whereby corporate applicants that have filed for patent protection in the recent past are more likely to be matched to a BvD company. This is plausible, given that companies 'inactive' in publishing annual accounts for more than five years have been discarded from the BvD financial database. Also, applicants associated with older patents are more likely to have been affected by merger and acquisition activity, name changes, defaults, etc.

Figure 3: Applicants and patents matched to BvD companies after both matching rounds, broken down by application filing year

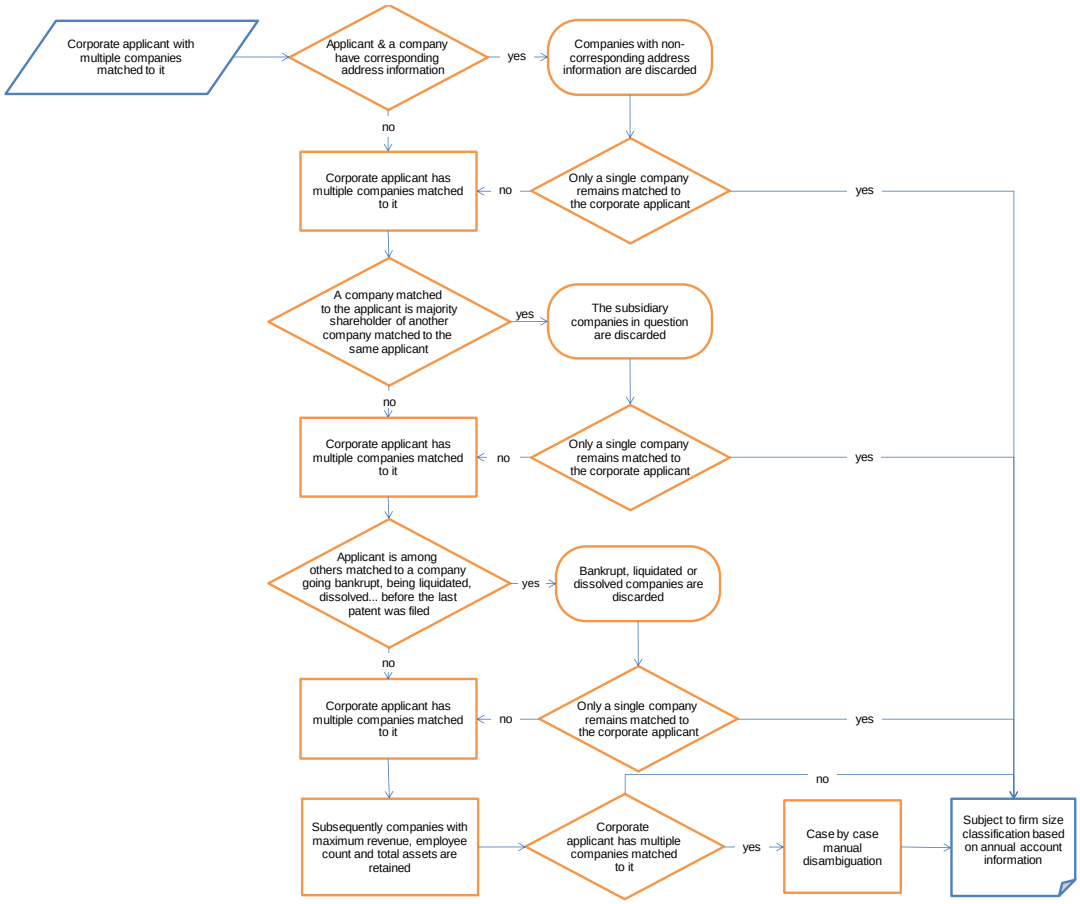


Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Disambiguation

While corporate law favours the idea that company names should be unique, the exact matching procedure explained above may lead to a single harmonised corporate applicant being linked to multiple harmonised BvD company names. A number of selection rules are applied to disambiguate these associations – the steps to identify the 'right' company are set out in **Figure 4**. The full disambiguation process consists of several rounds applied consecutively until a single match remains.

Figure 4: Process flow chart for disambiguation procedure



The disambiguation procedure involves the following steps:

i. companies with addresses that do not correspond to the address of the corporate applicant are removed if at least one other company is matched with the corresponding address information;

ii. priority is given to companies at the top of the shareholder hierarchy. When one of the matched companies holds the majority of shares in another company matched to the same corporate applicant, the latter is discarded;

iii. any liquidated, dissolved, bankrupt or inactive company matched with a corporate applicant that has filed patents since it was active (i.e. since the last year in which it filed annual accounts) is also discarded ([5]); and

iv. in line with Squicciarini & Dernis (2012), sequential rules are implemented in the final disambiguation rounds on the basis of firm size (records showing maximum values for revenue, staff count and total balance, in that order).

This procedure may lead to an underestimate of the proportion of SMEs among patenting companies, since matches for which size information is available will yield more final matches. If, however, one assumes that the majority of multiple matches involve companies belonging to the same business group, it makes sense to select the largest company among them, as this is the best indicator of group size. The

([5]) Disambiguation methodologies comparing the date of incorporation of matched corporate applicants with the date they filed their first patent were also explored. However, discarding companies with a negative patent lag (between year of first patent filed and year of incorporation) yielded too many sub-optimal matches. Over time, corporate restructuring may result in transfers of operations and assets to new legal entities. The last available accounts for the remaining 'shell' companies are not a reliable indicator of their current size.

limited number of multiple matches that remain after these automated disambiguation rounds (14 corporate applicants) are considered case by case.

To verify the accuracy of the matching and disambiguation methodology, we looked more closely at 50 % of the Belgian applicants, and 50 % of the Irish applicants, matched to a company in Amadeus 2012 from the same country, also examining official business registers with more detailed historical information on the establishment of domestic companies ([6]). In the case of Belgium, the findings revealed that 11 % of the matches were false positives (6 % in patent volume). For Ireland, 8 % of the patenting companies (8 % in patent volume) were incorrectly matched, i.e. to a non-corresponding corporate entity. The odds of false positives are higher among second-round matches. Computing the proportion of corporate applicants matched to SMEs may be affected by a certain degree of upward or downward bias (per country), yielding a theoretical over- or under-estimate. At the same time, we have no clear indication that the accuracy obtained is linked to the size of the company. This can therefore be regarded as 'noise' with no effect on further outcomes.

Classification

On the basis of European Commission Recommendation 2003/361/EC, a new SME definition was adopted on 1 January 2005, incorporating updated thresholds for companies applying for the European support programme for SMEs (see Figure 5). Our assessment of the firm size of matched companies is based primarily on this definition, which basically sets out five criteria for SME status: staff headcount (FTE), annual turnover, annual balance sheet total and previous criteria for partner and affiliated (controlling) enterprises. As shown in Figure 5, the firm must adhere to the staff headcount thresholds and either the turnover or the balance sheet ceiling.

Figure 5: SME headcount, annual turnover and annual balance sheet thresholds

Enterprise category	Headcount: Annual Work Unit (AWU)	Annual turnover		Annual balance sheet total
Medium-sized	< 250	≤ € 50 million	or	≤ € 43 million
Small	< 50	≤ € 10 million	or	≤ € 10 million
Micro	< 10	≤ € 2 million	or	≤ € 2 million

Source: European Commission, 2005.

With respect to a company's ownership/shareholder structure, the Commission defines three company types in order of increasing dependency: autonomous, partner and linked firms. Partly in line with previous research (Perrin & Speck, 2004; Belenzon & Berkovitz, 2007; Thoma *et al.*, 2010), we focused on 'linked' and 'non-linked' or 'independent' firms. 'Linked' enterprises form a group when direct or indirect control of the majority of voting rights results in a dominant influence on all enterprises involved,

([6]) Matches were verified using information from the *Kruispuntbank der Ondernemingen* for Belgium and the Company Registration Office database for Ireland.

in which case it is better to base firm-size classification on group level numbers ([7]). Direct or indirect ownership of at least 50 % of the shares suggests that one enterprise has a controlling position. The EC 2003 SME definition specifies that, to assess a firm's size, its financials should be fully consolidated with linked shareholders: linked companies form business groups and should be evaluated at that level. Unlike truly independent SMEs, small corporate entities fully owned by larger companies will benefit from their financial strength, managerial capacities and scale economies ([8]). More specifically, unlike their independent counterparts, 'linked' patenting SMEs may benefit from centralised R&D services and the intellectual property expertise at their disposal at business-group level.

Applying these financial and ownership criteria to a financial database is not straightforward. Financial databases covering all firm-size categories (rather than large or listed firms only) tend to suffer more from a lack of data for certain fields in the annual accounts. As will become apparent in the next section, BvD's Amadeus 2012 database is no exception in this respect.

Availability of financial and shareholder data

BvD has published new versions of Amadeus every year since it became a commercial product in 1996. It updates each version regularly throughout the year, incorporating newly published information. The firm-size assessment of patenting companies in this study is based on the most recent annual accounts available per firm. While time series data for the past 10 years are available for revenues, staff counts and total assets, ownership information is provided only for the last available financial year. This prevents us from dynamically adjusting the SME definition on the basis of shareholder information.

Table 2 shows a distribution of all Amadeus 2012 companies according to the last financial year for which BvD obtained annual accounts information, broken down by Member State.

([7]) For example, Amadeus categorises as 'dependent' small entities such as Tika Lakemedel AB (BvD ID SE5561300772) and Coley Pharmaceutical (BvD ID DE5050349817), which are controlled by pharmaceutical multinationals, Astra Zeneca and Pfizer. The French company Sogepass (BvD ID FR330649815) also complies with SME criteria (apart from dependence), but over 50 % of its shares are held by steel multinational ArcelorMittal SA.

([8]) We refer to small or large '(corporate) entities' when firm-size evaluation is based on financial size indicators only. To identify actual SMEs, we have to assess shareholding as well (see below).

Table 2: Financial years to which the most recent Amadeus 2012 company data refer (%)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	account info
EU-27	**1.60**	**0.49**	**0.59**	**0.78**	**1.39**	**1.84**	**2.84**	**4.65**	**6.07**	**15.29**	**57.69**	**6.31**	**0.46**	**100.00**
BE	10.79	1.65	1.62	1.61	1.70	1.57	1.73	1.97	2.28	3.77	56.86	4.03	10.42	100.00
BG	1.53	2.16	1.81	1.54	1.09	2.43	2.32	6.79	7.04	2.49	70.73	0.06	0.00	100.00
CZ	0.07	0.04	0.07	0.56	1.46	1.39	1.75	2.79	4.48	10.49	74.20	1.06	1.62	100.00
DK	0.00	0.00	0.00	0.00	0.00	0.00	2.11	4.36	4.95	4.94	60.33	23.32	0.00	100.00
DE	0.73	0.22	0.21	0.32	0.42	0.54	2.66	3.50	5.64	23.42	48.19	14.08	0.05	100.00
EE	0.00	1.91	2.25	2.45	2.03	1.92	2.50	3.97	5.59	7.68	69.27	0.42	0.00	100.00
IE	3.14	2.88	2.67	2.68	3.10	4.37	4.82	5.28	7.44	18.39	43.54	1.67	0.00	100.00
EL	0.00	0.00	0.37	0.47	0.50	0.85	1.13	2.26	4.44	9.49	77.52	2.98	0.00	100.00
ES	5.49	1.46	2.40	3.45	3.32	3.56	4.74	8.06	7.34	18.05	42.14	0.00	0.00	100.00
FR	0.00	0.01	0.24	0.32	0.44	0.59	0.91	2.20	7.73	14.67	67.45	5.44	0.00	100.00
IT	0.23	0.10	0.24	0.32	0.41	0.37	0.42	5.67	6.65	10.18	74.79	0.62	0.00	100.00
CY	0.00	0.00	0.00	0.00	0.00	1.27	1.88	6.35	74.59	7.51	6.29	2.11	0.00	100.00
LV	1.10	0.23	0.25	0.12	1.64	1.09	2.74	8.27	9.69	13.41	59.85	1.61	0.00	100.00
LT	0.08	0.05	0.03	0.11	0.08	0.02	0.00	6.99	10.96	11.93	69.71	0.04	0.00	100.00
LU	1.08	0.92	1.61	3.33	2.08	3.91	6.62	5.72	9.59	29.70	30.88	4.57	0.00	100.00
HU	0.02	0.00	0.00	0.00	0.03	3.43	0.17	0.16	1.41	8.68	86.08	0.00	0.00	100.00
MT	1.36	0.51	0.43	0.63	0.73	1.37	2.02	2.25	11.10	66.71	12.87	0.02	0.00	100.00
NL	7.12	1.68	1.41	1.26	1.27	2.06	2.59	3.13	4.99	25.26	48.60	0.62	0.00	100.00
AT	0.19	0.04	0.03	0.02	0.07	0.46	3.09	3.64	11.34	10.76	42.40	27.96	0.00	100.00
PL	0.10	0.03	0.04	0.04	0.06	0.06	0.47	3.39	3.87	67.96	23.98	0.00	0.00	100.00
PT	0.00	0.00	0.00	0.00	0.00	0.65	4.52	6.17	7.39	11.19	70.07	0.00	0.00	100.00
RO	0.04	0.03	0.05	0.10	0.17	0.54	0.89	2.16	4.90	5.75	85.35	0.00	0.00	100.00
SI	0.00	0.00	0.00	0.00	0.03	0.12	0.52	0.88	1.08	1.39	2.01	93.96	0.00	100.00
SK	0.07	0.06	0.08	0.14	20.32	2.73	4.03	5.87	18.71	7.73	40.25	0.01	0.00	100.00
FI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.81	8.65	72.99	17.53	0.00	100.00
SE	0.32	0.20	0.07	0.33	0.29	0.41	0.40	1.18	1.41	4.18	77.15	14.07	0.00	100.00
UK	0.52	0.21	0.32	0.67	1.99	4.24	6.49	7.62	6.39	7.11	54.31	10.14	0.00	100.00

Note: A colour scale applies for percentages, ranging between the lightest shade of white for minima and the darkest shade of orange for maxima.

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Although the greater part of the most recent financial information dates back to financial year 2010, this truncation allows for a degree of fit with applicants in the October 2011 edition of PATSTAT: given the publication delay of at least 18 months for USPTO and EPO patents, the most recent corporate applicants in PATSTAT are likely to be companies operating at the end of 2010/beginning of 2011.

To evaluate the extent to which the EC 2003 SME definition can be used to differentiate between SMEs and large companies in Amadeus 2012, we assessed per-country data availability for the indicators of interest. **Table 3** shows coverage rates per indicator in Amadeus 2012.

Table 3: Amadeus – overall coverage of financial and ownership indicators required to determine firm size (per indicator)

	Total number of companies	Companies reporting operational revenues		Companies reporting staff count		Companies reporting total assets		Companies reporting dependency status	
		number	%	number	%	number	%	number	%
BE	609412	129201	21.20	257647	42.28	545187	89.46	41646	6.83
BG	494532	59231	11.98	488209	98.72	60490	12.23	336347	68.01
CZ	482679	469799	97.33	294733	61.06	184120	38.15	198742	41.17
DK	25230	42447	16.78	76887	30.40	252918	100.00	144339	57.07
DE	1456074	419446	28.81	437928	30.08	1101434	75.64	1170990	80.42
EE	108986	94667	86.86	52709	48.36	108936	99.95	84675	77.69
IE	211372	25951	12.28	21836	10.33	199798	94.52	16627	7.87
EL	28401	28401	100.00	23600	83.10	28401	100.00	21289	74.96
ES	1273351	1140063	89.53	843380	66.23	1273351	100.00	355419	27.91
FR	1291883	1291875	100.00	878954	68.04	1291882	100.00	269729	20.88
IT	119884	1188353	99.14	352781	29.43	1198684	100.00	856313	71.44
CY	41289	907	2.20	1878	4.55	1005	2.43	36963	89.52
LV	110292	85711	77.71	102382	92.83	7938	7.20	5406	4.90
LT	117370	26789	22.82	110033	93.75	3712	3.16	9636	8.21
LU	19240	5199	27.02	4040	21.00	16028	83.31	15838	82.32
HU	377912	316267	83.69	135538	35.86	375792	99.44	12683	3.36
MT	15259	15259	100.00	392	2.57	15259	100.00	4907	32.16
NL	895494	31108	3.47	643147	71.82	822005	91.79	272414	30.42
AT	224480	6385	2.84	167413	74.58	77880	34.69	140733	62.69
PL	960971	117796	12.26	902969	93.96	121316	12.62	589289	61.32
PT	434526	365782	84.18	343776	79.12	428069	98.51	319629	73.56
RO	571289	568039	99.43	566221	99.11	571038	99.96	493193	86.33
SI	76089	2512	3.30	10558	13.88	2574	3.38	15661	20.58
SK	230781	165399	71.67	197956	85.78	58557	25.37	17996	7.80
FI	170484	160798	94.32	47776	28.02	170484	100.00	24134	14.16
SE	866641	829664	95.73	848854	97.95	319732	36.89	67869	7.83
UK	3076136	447311	14.54	132348	4.30	2998120	97.46	1918280	62.36
Total	**15596557**	**8034360**	**51.51**	**7943945**	**50.93**	**12234710**	**78.44**	**7440747**	**47.71**

Note: a colour scale applies for percentages, ranging between the lightest shade of white for minima and the darkest shade of orange for maxima.

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

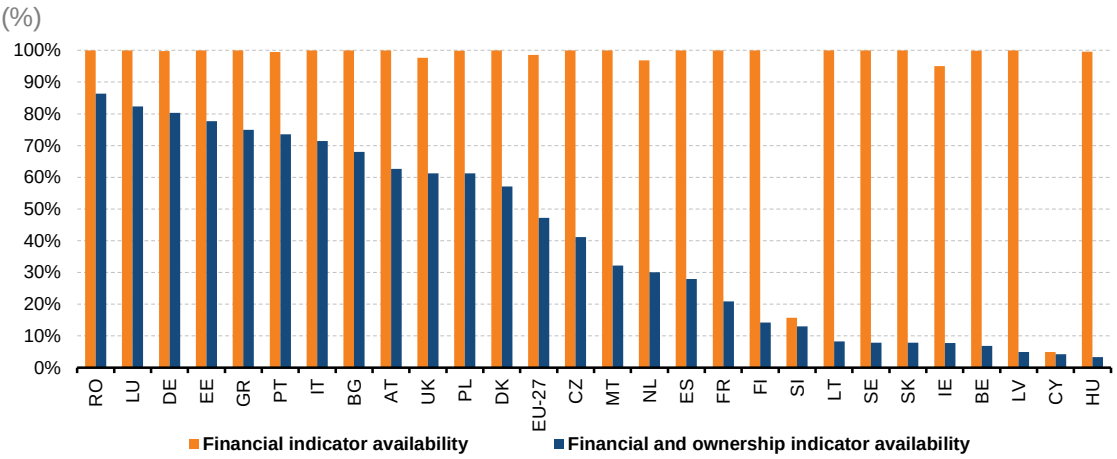
Table 3 shows non-trivial data gaps in Amadeus 2012, with considerable variation across Member States. This is caused by several factors:

- Under some national legislation, certain types of company (e.g. those below certain size thresholds or with simple ownership structures) are not bound to full disclosure and may publish simplified annual accounts or be exempt from disclosing financial information altogether; and
- the financial database evaluated does not necessarily comprise all publicly available financial information (some firms, or specific information on firms, missing).

Macartney (2007) reported that the number of companies covered by BvD (Amadeus) had grown significantly from 2004 onwards. Amadeus 2012 covers twice as many companies as Amadeus 2007, suggesting that coverage is improving and that the increase is not simply due to more companies being established ([9]).

([9]) Of the 15596557 companies (EU-27) in Amadeus (excluding the 1046637 French companies for which additional credits were needed), only 3806366 were established after 2006 and could not have been part of Amadeus 2007. Consequently, over 50 % of the extra coverage in Amadeus 2012, as compared with Amadeus 2007, cannot be explained by new company establishments.

Figure 6: Amadeus – overall coverage of (financial and financial/ownership) indicators required to determine firm size



Source: Amadeus 2012

Figure 6 provides further insight into the availability of firm-size indicators in Amadeus 2012. The Member States are ranked according to descending rates of joint availability of financial and ownership indicators. Overall, 99 % of the firms in Amadeus 2012 report at least one of the three financial size indicators from which firm size can be derived independently of ownership information ([10]). Cyprus and Slovenia appear to be the only 'problematic' cases in terms of coverage. The figures are not so good if one takes ownership information availability into account. Only 47 % of the firms provide at least one of the financial firm-size indicators and sufficient shareholder information to distinguish dependent from independent firms. In the next section, we elaborate on how companies with limited firm-size information are assigned to a firm-size category.

First round of classification: using financial size indicators to differentiate between large and small entities

We classify matched corporate applicants in two stages:

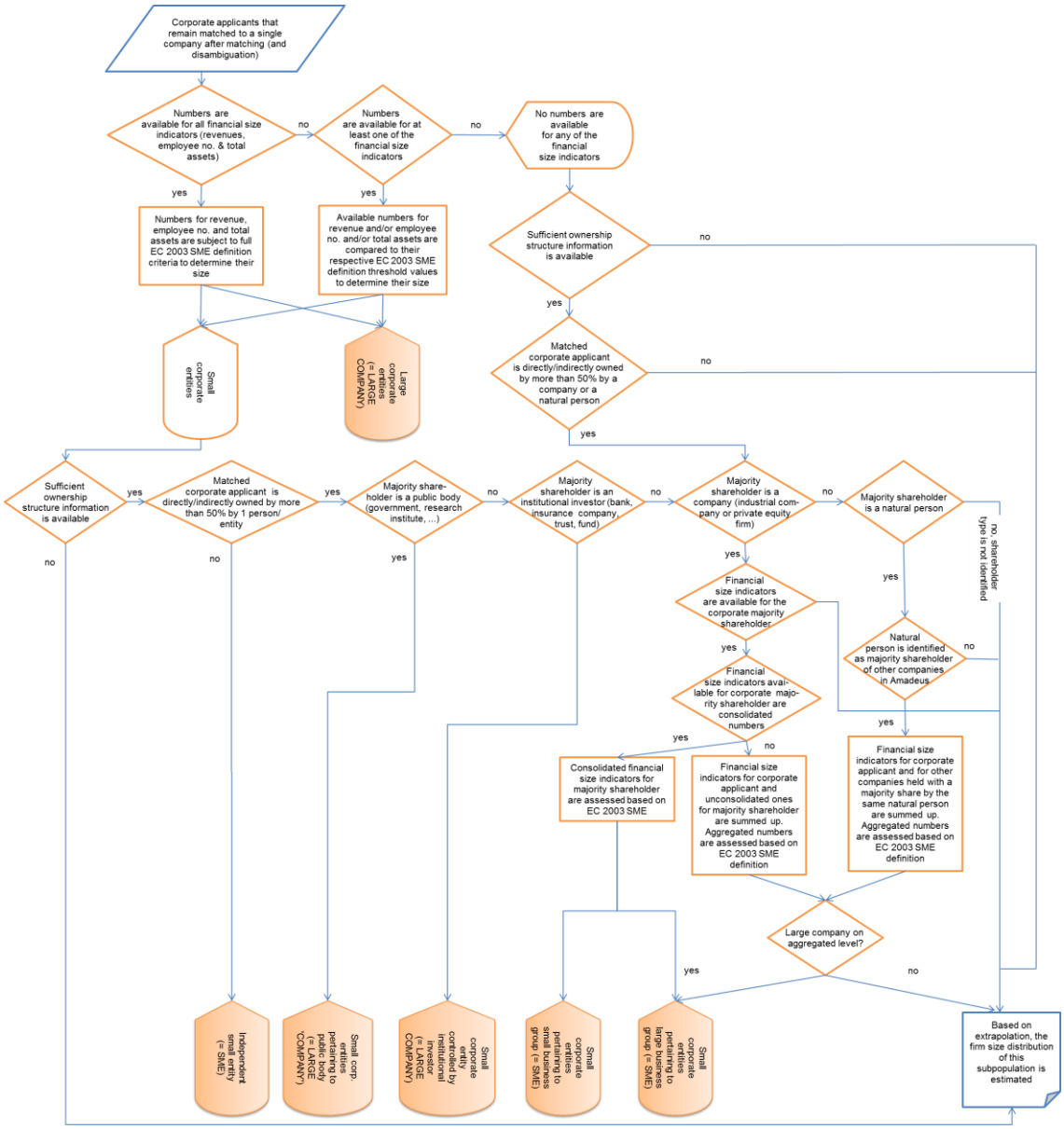
i. firm-size indicators only (revenues, employee count and total assets) are used to distinguish between large and small corporate entities; and

ii. to filter out the *actual* SMEs among the small corporate entities, shareholder information is introduced and consolidated financial size indicators are assessed in the case of majority corporate ownership.

A process flow chart showing the classification procedure step by step is presented in **Figure 7**.

([10]) Helmers and Rogers (2009), for instance, assume that companies for which no financials are reported are micro firms.

Figure 7: Process flow chart for classification procedure



We used the most recent available indicators in Amadeus 2012 to identify the size of corporate entities (see above). A first approach to distinguishing between large and small corporate entities adheres strictly to EC directives on annual revenues, staff count and balance sheet total. According to this baseline definition, entities that are 'certainly large' have either:

- revenues over EUR 50 million and total assets over EUR 43 million; or
- a staff count of 250 FTEs or more.

Companies that appear as small entities report:

- a staff count of less than 250; and
- revenues of EUR 50 million or less or total assets of EUR 43 million or less.

The limited availability of some indicators (see **Table 3**) underscores the need for a second classification round, using looser, mutually exclusive definitions for companies that are 'most likely' small or large entities. Missing financial information is extrapolated from the limited available data. If, for instance, one available indicator lies above the threshold, the other indicators are assumed to lie above their thresholds as well. According to this approach, companies that are 'most likely' large are those reporting:

- revenues over EUR 50 million, where staff count and total asset numbers are unavailable;
- total assets over EUR 43 million, where staff count and revenue numbers are unavailable;
- revenues over EUR 50 million and total assets of EUR 43 million or less, where staff count numbers are unavailable; and
- revenues of EUR 50 million or less and total assets over EUR 43 million, where staff count numbers are unavailable.

Companies that are 'most likely' small entities are those reporting:

- a staff count of less than 250 FTEs, where revenues and total assets are unavailable;
- revenues of EUR 50 million or less and total assets of EUR 43 million or less, where staff count numbers are unavailable;
- total assets of EUR 43 million or less, where staff count and revenue numbers are unavailable;
- revenues of EUR 50 million or less, where staff count and total assets numbers are unavailable;
- revenues of EUR 50 million or less and total assets of EUR 43 million or less, where staff count numbers are unavailable;
- revenues of over EUR 50 million and staff counts of less than 250 FTEs, where total assets numbers are unavailable; and
- total assets of over EUR 50 million and staff counts of less than 250 FTEs, where revenue numbers are unavailable.

Obviously, caution is called for in performing this kind of extrapolation and interpreting indicators based on them, since the validity of the underlying assumptions remains unclear. In addition, as stated above, ownership information should be taken into account to determine whether small entities are actually SMEs. The following section addresses the methodological challenges that we met in incorporating ownership information into the firm-size assessment.

Second round of classification: using ownership information to identify actual SMEs

The shareholder information in Amadeus is essential for determining which companies truly qualify as independent SMEs, which are members of larger (multinational) business groups and which are backed by other types of shareholders such as governments, institutional investors or universities. To help users identify independent companies, BvD has created an 'independence indicator' showing how independent a company is *vis-à-vis* its shareholders. We use this indicator to determine which of the matched companies require further exploration as regards their shareholder structure.

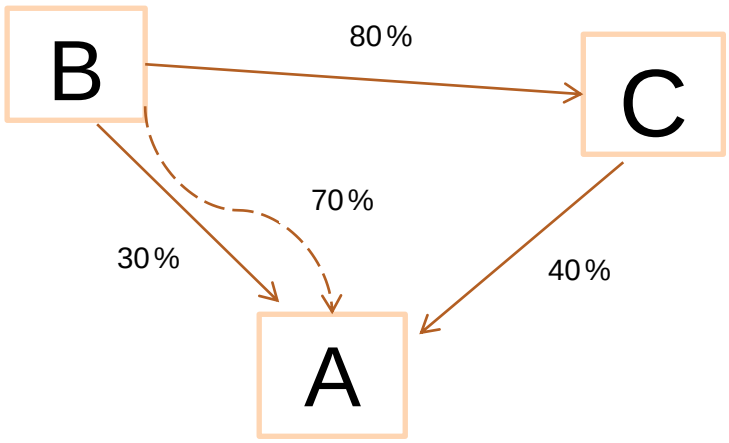
This further exploration is based on actual share percentages per shareholder. BvD registers two measures per shareholder:

- a direct share percentage, i.e. the percentage of the company directly owned by the shareholder; and
- a total share percentage, i.e. the sum of direct ownership percentages and indirect ownership percentages – where the shareholder holds (additional) company shares through (other) subsidiaries.

Where the ultimate owner controls the intermediate subsidiary by owning a majority of its shares, one can

assume that its ownership stake in the company under consideration is a reflection of the share percentage its subsidiary holds in that company. Otherwise, BvD calculates the indirect percentage by multiplying the ultimate owner's direct share in the intermediate subsidiary by the direct share that the subsidiary holds in the company under consideration.

Figure 8: Illustration of BvD total ownership percentage calculation



Source: BvDEP Ownership Database, 2008.

Figure 8 provides an illustration of a shareholder structure that is examined in this way. Company A is directly owned by Company B (30 %) and Company C (40 %). Company C is also directly owned by B (80 %). As B controls C, the calculated total percentage between A and B is 30 % + 40 % = 70 %. More specific information on BvD's procedures for calculating total ownership percentages can be found in the BvD *Amadeus Ownership Guide* (2008). As we are mainly interested in the economic owner, total ownership percentages (if available) are preferred to direct ownership percentages. The latter are used only if the former are unavailable. Where this information is available, the size of corporate applicants controlled by business groups is evaluated at group level (see below).

Results: small vs. large corporate entities and SMEs vs. large companies

The results of the first round, whereby EU corporate applicants are classified as small or large corporate entities, are presented in Annex 2 (**Table 18**). The outcomes show that 8.9 % qualify as large corporate entities in the first round, whereas 52.5 % can be characterised as small corporate entities. 8.5 % of the large entities are identified as 'certainly' and the remaining 0.4 % as 'most likely' large corporate entities (see above). 27.2 % of the small entities are characterised as 'certainly' and 25.3 % are 'most likely' small (see above). The corresponding patent volumes are presented in **Table 4** ([11]), the first row of which contains the total number of distinct patents per size category for the EU overall.

([11]) Patent applications filed by multiple corporate co-applicants from the same country are counted multiple times according to the number of co-applicants sharing the same nationality. This has a limited impact on the results as compared with counting such applications only once: the percentage contribution to patenting remains stable. Patents filed by co-applicants from different Member States are counted more than once at country level, according to the number of countries to which the co-application is assigned.

Table 4: Results of first round of classification of corporate applications based on financial firm-size indicators only

Country	Large				Small				No financial size indicators		Not matched		Total
	Certainly		Most likely		Certainly		Most likely						
	#	%	#	%	#	%	#	%	#	%	#	%	#
EU-27	696 716	52.9	25 896	2.0	174 998	13.3	192 480	14.6	4 259	0.3	222 219	16.9	1 316 568
BE	15 103	57.8	1 633	6.2	5 301	20.3	1 182	4.5	1	0	2 909	11.1	26 129
BG	37	19.5		0	19	10	17	8.9		0	117	61.6	190
CZ	362	25	5	0.3	518	35.7	82	5.7		0	483	33.3	1 450
DK	15 229	51.6	94	0.3	5 821	19.7	3 292	11.2	32	0.1	5 019	17	29 487
DE	293 215	54.5	7 412	1.4	35 956	6.7	116 945	21.7	218	0	84 101	15.6	537 847
EE	6	2.7	2	0.9	92	40.7	36	15.9		0	90	39.8	226
IE	1 225	14	315	3.6	1 472	16.8	3 434	39.2	129	1.5	2 192	25	8 767
EL	29	4.3		0	98	14.5	64	9.5	5	0.7	480	71	676
ES	4 838	28.4	18	0.1	5 900	34.7	738	4.3		0	5 525	32.5	17 019
FR	107 682	60	4 473	2.5	26 879	15	5 031	2.8	47	0	35 345	19.7	179 457
IT	30 752	39.8	343	0.4	23 066	29.9	6 192	8	5	0	16 828	21.8	77 186
CY	6	0.6	4	0.4	66	7.1	45	4.8	202	21.7	609	65.3	932
LV	9	3.1		0	18	6.3	10	3.5		0	251	87.2	288
LT		0		0	6	22.2	7	25.9		0	14	51.9	27
LU	1 199	22.2	103	1.9	249	4.6	1 556	28.8		0	2 292	42.5	5 399
HU	310	18.4		0	229	13.6	95	5.6	2	0.1	1 053	62.3	1 689
MT	56	13.1		0	18	4.2	289	67.8		0	63	14.8	426
NL	89 064	67	7 173	5.4	16 027	12.1	7 474	5.6	1 577	1.2	11 550	8.7	132 865
AT	11 072	43.8	26	0.1	1 447	5.7	6 042	23.9	1	0	6 705	26.5	25 293
PL	271	23	17	1.4	194	16.5	314	26.6		0	383	32.5	1 179
PT	223	20.9	9	0.8	434	40.8	72	6.8		0	327	30.7	1 065
RO	5	5.3		0	29	30.5		0		0	61	64.2	95
SI	116	8.1		0	35	2.4	102	7.1	425	29.6	760	52.9	1 438
SK	59	19.3		0	132	43.3	34	11.1		0	80	26.2	305
FI	35 366	69.3	1 756	3.4	4 260	8.3	3 452	6.8	40	0.1	6 178	12.1	51 052
SE	34 513	40.7	356	0.4	16 987	20	1 193	1.4	32	0	31 763	37.4	84 844
UK	55 969	42.6	2 157	1.6	29 745	22.7	34 782	26.5	1 543	1.2	7 041	5.4	131 237

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Table 4 shows that 8.9% of large corporate applicants hold over 54.9% of patents, whereas 52.5% of small entities account for only 27.9% of the total matched patent volume.

In the second round, shareholder information for possible SMEs (i.e. small corporate entities and those companies lacking financial size indicators altogether) is taken into consideration to identify actual SME activity. Further dependency information is deemed irrelevant for companies already identified as large corporate entities. Among the small entities, independent SMEs are those companies that have no majority corporate shareholders (as indicated in Amadeus) ([12]). For a few (0.5%) of the remaining

([12]) These counts should still be treated with caution, since they include companies complying with both the strict 'certain' SME definition and the looser 'most likely' SME definition.

possible (non-independent) SMEs, no financial size indicators are available. Most of these are companies that qualify as dependent small entities (26.4 % of corporate applicants) or small entities with an unknown degree of independence (15.0 %). However, of the companies in the last three categories, we were able to reassign some to the large company category on the basis of information about their majority shareholders. **Table 5** presents the distribution of these majority shareholders according to sector.

Table 5: Dependent, patenting SMEs by majority shareholder type

Country	Industrial company		Institutional investor		Natural person		Private equity firm		Public body		Total
	#	%	#	%	#	%	#	%	#	%	
EU-27	16564	59.9	639	2.3	10133	36.6	78	0.3	249	0.9	27663
BE	338	90.9	10	2.7	20	5.4	4	1.1	-	0.0	372
BG	4	33.3	1	8.3	6	50.0	1	8.3	-	0.0	12
CZ	39	29.5	-	0.0	92	69.7	1	0.8	-	0.0	132
DK	736	81.3	67	7.4	96	10.6	4	0.4	2	0.2	905
DE	3886	49.3	65	0.8	3887	49.3	13	0.2	35	0.4	7886
EE	13	54.2	2	8.3	9	37.5	-	0.0	-	0.0	24
IE	180	91.4	4	2.0	13	6.6	-	0.0	-	0.0	197
EL	6	24.0	-	0.0	19	76.0	-	0.0	-	0.0	25
ES	400	49.6	27	3.3	375	46.5	3	0.4	2	0.2	807
FR	1435	71.8	122	6.1	432	21.6	8	0.4	3	0.2	2000
IT	1807	44.8	46	1.1	2172	53.8	6	0.1	6	0.1	4037
CY	12	50.0	1	4.2	11	45.8	-	0.0	-	0.0	24
LV	3	75.0	-	0.0	1	25.0	-	0.0	-	0.0	4
LT	3	100.0	-	0.0	-	0.0	-	0.0	-	0.0	3
LU	67	62.6	-	0.0	40	37.4	-	0.0	-	0.0	107
HU	10	76.9	-	0.0	3	23.1	-	0.0	-	0.0	13
MT	16	94.1	-	0.0	1	5.9	-	0.0	-	0.0	17
NL	1757	87.8	91	4.5	35	1.7	2	0.1	117	5.8	2002
AT	568	59.5	12	1.3	324	34.0	2	0.2	48	5.0	954
PL	33	40.2	1	1.2	48	58.5	-	0.0	-	0.0	82
PT	35	54.7	1	1.6	26	40.6	2	3.1	-	0.0	64
RO	3	30.0	-	0.0	7	70.0	-	0.0	-	0.0	10
SI	22	43.1	1	2.0	28	54.9	-	0.0	-	0.0	51
SK	9	60.0	-	0.0	6	40.0	-	0.0	-	0.0	15
FI	288	69.6	11	2.7	108	26.1	5	1.2	2	0.5	414
SE	1045	89.7	41	3.5	67	5.8	10	0.9	2	0.2	1165
UK	3849	60.7	136	2.1	2307	36.4	17	0.3	32	0.5	6341

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Of the possible SMEs that are non-independent (dependent small entities, small entities with unknown dependency status and companies lacking any financial firm-size indicator), 0.9 % are owned by public bodies, including (local) authorities, research institutions, foundations and universities. In 36.6 % of the remaining possible SMEs, a natural person (generally a family or an individual) holds the majority of the shares. 61.1 % are controlled by corporate players, most of which are industrial companies (59.9 %),

followed by institutional investors (2.3 %) and private equity companies (0.3 %).

In line with EU directives, the financials of companies that belong to larger business groups are assessed on a consolidated basis. Companies in Amadeus can be members of two types of business group. Companies controlled by another company constitute a formal business group, while those controlled by a natural person who also holds majority stakes in other companies are part of an informal group.

Of the 43 631 potential SMEs that are non-independent (see Annex 2 – Table 19), 16 755 have a corporate organisation ([13]) as majority shareholder (holding 50+ % of the shares). 13 920 of these are companies established in a Member State and are not 'financial institutions' or investment entities, so we can associate them with their individual Amadeus record ([14]). For 5 030 (or 36 %) of the 13 920 EU majority shareholders, the Amadeus indicators refer to the consolidated level. Of all corporate applicants, 1.9 % (accounting for 1.1 % of corporate applications) are part of a small business group according to the mother company's consolidated financials, and 2.8 % (3.9 % of corporate applications) belong to a large business group.

For the non-independent potential SMEs with unconsolidated mother company financials (e.g. non EU majority shareholders), the consolidation methodology proposed in the EC 2003 SME definition is also applied. Consolidated group-level figures are approximated by adding the revenue, staff count and total asset figures of the majority shareholder(s) to the possible SMEs' financials. This assumes that all parent company financials provided by BvD are unconsolidated unless otherwise specified. On the basis of this cruder consolidation strategy, an additional 2.5 % of corporate applicants (2.4 % of corporate applications) can be assumed to belong to a large business group and so do not qualify as true SMEs.

A similar approach was taken for companies with natural persons as majority shareholders. If Amadeus registered these natural persons as holding majority stakes in other companies, the financials across all majority-owned companies were added together and compared with large-company thresholds. On the basis of this cruder consolidation strategy, the use of shareholder information resulted in an additional 0.1 % of corporate applicants (0.1 % of corporate applications) being categorised as large companies, rather than true SMEs.

By aggregating the results from all consolidation approaches, we can reclassify 5.4 % of the 41.9 % non-independent possible SMEs among the corporate applicants as large companies. This represents 6.4 % of the matched patent volume. The 1.9 % of possible SMEs assigned to the small business group category account for 1.1 % of the matched patent volume. This is an underestimate of the number of companies in the three categories of possible SMEs that can be reallocated to the large and small company categories, as the BvD ownership structure data are incomplete and the same constraints apply as regards the availability of financial information for corporate shareholders.

The classification procedure results in 10 categories of corporate applicants and applications, as presented in **Table 6** (applicant-level results are presented in **Table 20** in Annex 2).

([13]) i.e. industrial companies, holding companies and private equity firms. Majority shareholders in the form of institutional investors such as pension and mutual funds/trusts, banks and insurance companies are treated separately.

([14]) For non-European shareholders, the only relevant information in Amadeus is turnover, total assets and staff count. Only EU-27 information was downloaded from Amadeus, so more detailed annual accounts information for companies with European, non-EU (e.g. Swiss) majority shareholders was not considered (although available in Amadeus).

Table 6: Overall matching and size classification of corporate applications
(%)

Country	Large	Small — large group	Small — maj. owned by institutional investor	Small — maj. owned by public body	Small — small group	Small and independent	Small — maj. owned by natural person	Small with insufficient ownership information	Matched but insuff. fin. and ownership information	Not matched	Total
EU-27	54.9	6.4	0.3	0.1	1.1	4.3	5.5	10.2	0.2	16.9	100.0
BE	64.1	7.3	0.1	0.0	0.3	4.3	0.3	12.7	0.0	11.1	100.0
BG	19.5	2.1	0.0	0.0	0.0	3.2	6.3	7.4	0.0	61.6	100.0
CZ	25.3	3.9	0.0	0.0	0.5	7.9	11.2	17.9	0.0	33.3	100.0
DK	52.0	3.8	0.8	0.0	2.9	10.9	1.0	11.6	0.1	17.0	100.0
DE	55.9	7.7	0.1	0.0	0.2	3.0	9.7	7.7	0.0	15.6	100.0
EE	3.5	2.2	2.2	0.0	0.0	24.3	10.2	17.7	0.0	39.8	100.0
IE	17.6	11.8	0.1	0.0	1.9	2.0	0.7	39.5	1.5	25.0	100.0
EL	4.3	1.0	0.0	0.0	0.0	6.1	11.1	5.8	0.7	71.0	100.0
ES	28.5	5.6	1.6	0.0	1.2	5.9	5.5	19.3	0.0	32.5	100.0
FR	62.5	4.4	0.8	0.0	0.2	2.9	1.2	8.2	0.0	19.7	100.0
IT	40.3	4.3	0.2	0.0	1.0	13.0	9.2	10.2	0.0	21.8	100.0
CY	1.1	2.6	0.0	0.0	0.2	0.8	0.5	8.0	21.5	65.3	100.0
LV	3.1	1.0	0.0	0.0	0.0	2.4	0.3	5.9	0.0	87.2	100.0
LT	0.0	7.4	0.0	0.0	0.0	7.4	0.0	33.3	0.0	51.9	100.0
LU	24.1	1.4	0.0	0.0	0.0	2.1	2.2	27.6	0.0	42.5	100.0
HU	18.4	0.1	0.0	0.0	0.0	0.9	0.2	18.1	0.1	62.3	100.0
MT	13.1	0.7	0.2	0.0	1.6	6.8	0.7	62.0	0.0	14.8	100.0
NL	72.4	5.0	0.2	0.3	1.3	0.8	0.1	11.0	0.3	8.7	100.0
AT	43.9	5.3	0.3	1.4	0.0	6.4	5.1	11.1	0.0	26.5	100.0
PL	24.4	7.9	0.2	0.0	0.3	17.8	7.8	9.2	0.0	32.5	100.0
PT	21.8	6.6	0.0	0.0	1.1	19.8	5.9	14.1	0.0	30.7	100.0
RO	5.3	2.1	0.0	0.0	0.0	1.1	12.6	14.7	0.0	64.2	100.0
SI	8.1	0.7	0.1	0.0	0.0	2.6	2.5	4.2	29.1	52.9	100.0
SK	19.3	5.9	0.0	0.0	0.0	8.2	12.1	28.2	0.0	26.2	100.0
FI	72.7	1.8	0.0	0.0	0.9	2.3	0.8	9.3	0.1	12.1	100.0
SE	41.1	3.7	0.2	0.0	3.4	2.7	0.5	11.0	0.0	37.4	100.0
UK	44.3	11.3	0.6	0.1	4.7	10.2	5.5	16.9	1.0	5.4	100.0

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Table 21 in Annex 2 (for the applicant-level outcomes) complements **Table 6** by providing a more aggregated picture grouping corporate applicants in four classes: 'large' companies, actual SMEs, companies for which insufficient information is available for us to classify them as either large or small, and corporate applicants not matched to any company in the Amadeus directory. Each of the 10 'categories' is mapped to one of these four classes. Multiple categories are linked to the first three classes. The 'large entities' and SMEs that belong to large business groups are unquestionably 'large' companies. The small proportion of companies backed by (semi-)public actors are excluded from the category of actual SMEs and assigned to the 'large' company category. Similarly, companies controlled by institutional investors are classified as 'large' companies. The 'actual SMEs' consist of independent

SMEs and SMEs linked to small business groups. The class of companies for which information is insufficient for reliable firm-size determination comprises companies for which ownership (and financial) information is lacking, but also possible SMEs controlled by natural persons. The incompleteness of Amadeus and the possibility that those natural persons hold majority stakes in other companies prevent us from classifying these applicants directly as actual SMEs. To produce reliable indicators for all corporate applicants, additional efforts are required to estimate the proportion of SMEs in the last two classes, as we lack information to determine the size of the company or business group.

Table 7: Overall matching and 'large' vs. SME classification results for applications filed by corporate applicants

(%)

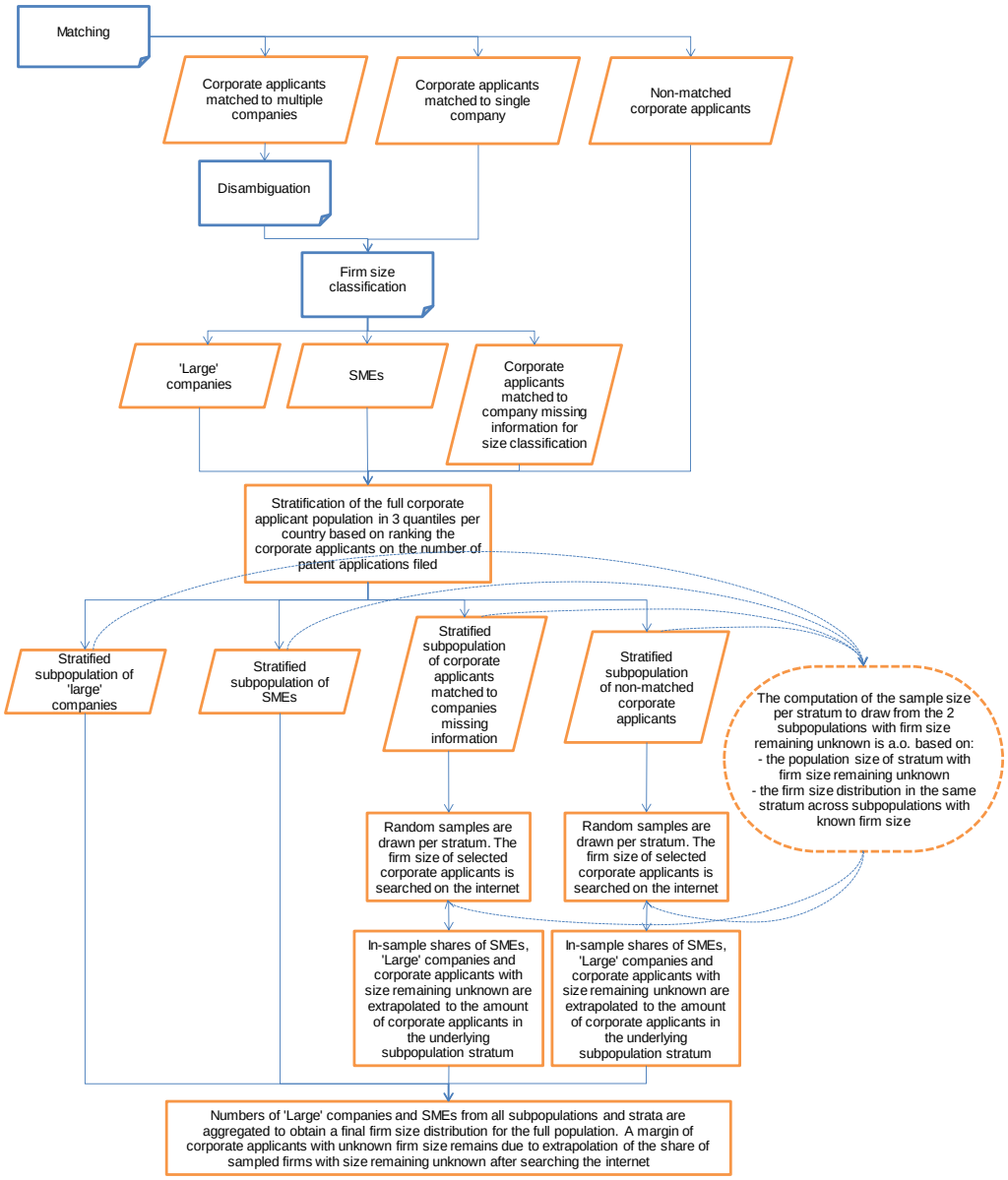
Country	'Large' company	SME	Matched but unknown	Not matched	Total
EU-27	61.7	5.4	16.0	16.9	100.0
BE	71.4	4.6	12.9	11.1	100.0
BG	21.6	3.2	13.7	61.6	100.0
CZ	29.2	8.4	29.1	33.3	100.0
DK	56.5	13.7	12.7	17.0	100.0
DE	63.8	3.2	17.5	15.6	100.0
EE	8.0	24.3	27.9	39.8	100.0
IE	29.5	3.8	41.7	25.0	100.0
EL	5.3	6.1	17.6	71.0	100.0
ES	35.7	7.1	24.7	32.5	100.0
FR	67.7	3.1	9.5	19.7	100.0
IT	44.8	14.0	19.4	21.8	100.0
CY	3.6	1.0	30.0	65.3	100.0
LV	4.2	2.4	6.3	87.2	100.0
LT	7.4	7.4	33.3	51.9	100.0
LU	25.5	2.2	29.8	42.5	100.0
HU	18.4	0.9	18.4	62.3	100.0
MT	14.1	8.5	62.7	14.8	100.0
NL	77.9	2.0	11.4	8.7	100.0
AT	50.9	6.5	16.2	26.5	100.0
PL	32.5	18.1	17.0	32.5	100.0
PT	28.4	20.9	20.0	30.7	100.0
RO	7.4	1.1	27.4	64.2	100.0
SI	8.8	2.6	35.7	52.9	100.0
SK	25.2	8.2	40.3	26.2	100.0
FI	74.5	3.2	10.2	12.1	100.0
SE	45.0	6.1	11.5	37.4	100.0
UK	56.3	14.9	23.4	5.4	100.0

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Extrapolation

As the previous findings produce a sizeable number of undecided cases (and hence unallocated patent volume), additional efforts are required to produce a more precise indicator of SMEs' contribution. This involves case-by-case searches for financial and ownership information which would allow us to classify firms as large or small. Given the size of the population involved, we propose taking a sample and extrapolating the findings to the (stratified) population.

Figure 9: Process flow chart of extrapolation procedure



Sample size computation methodology

We used the Eurostat 2012 Community Innovation Survey methodological recommendations (these build on Cochran [1977]) to determine the required sizes of the samples to be drawn from the target population of applicants not classified as large companies or SMEs. The sizes set were such as to allow us to make statements at a precision level of 5 % (95 % confidence intervals) on the overall proportion of SMEs in the target population. In order to limit sampling error, the target population was broken down into similarly sized subgroups or strata. The number of applications filed by an applicant is one of the few size-related indicators available in the patent database: companies filing many applications are more likely to be large companies. Therefore, the stratification into three quantiles of applicants per country was based on the applicants' contribution to the country's patent volume (33 %; 66 %; 100 %):

- first quantile: the most intensive applicants, accounting for 33 % of total patent volume;
- second quantile: applicants responsible for the next 33 % of applications; and
- third quantile: the remaining applicants.

To control for specificities of matched vs. unmatched applicants, a second level of stratification was added by compiling separate representative samples for:

a) 'matched but missing information' company names, i.e. previous, automated assessments match the company to the financial database but the information available (in the financial database) is insufficient to classify it as an SME or otherwise. Missing information can pertain to size (turnover, employment, total assets) and/or the structure of the group; and

b) 'unmatched' company names, i.e. automated matching with existing financial company databases yields no corresponding entities. This category was stratified explicitly, since it may be assumed that it contains smaller entities not bound to full financial disclosure.

The stratification procedure therefore breaks the population down into 27 (countries) * 3 (patent-volume quantiles) * 2 (matched but unknown size vs unmatched) strata, resulting in a total of 162 strata.

The sample size is calculated by means of stratified random sampling. The sample size n_D in the full target population D is calculated as:

$$n_D = \frac{(\sum_{h=1}^{H} W_h . S_h)^2}{\mathrm{V}(\hat{\theta}_D) + \frac{1}{N_D}\sum_{h=1}^{H} W_h . S_h^2}$$

where $\mathrm{V}(\hat{\theta}_D)$ is the variance of the estimated overall proportion of SMEs, H is the number of strata in the target population D, $W_h = N_h / N_D$, where N_h is the number of enterprises in stratum h; N_D is the total number of enterprises in target population D; and S_h^2 is the stratum variance for the SME dummy variable y_a.

The numerator reflects the variance within each stratum multiplied by the size of the stratum. In other words, for strata with more firms, more n will be included. A similar observation can be made with respect to variance (S_h): this will be highest when the proportion of large and small firms is equal (50 x 50 = 2500, whereas 90 x 10 = 900). For further technical details on the approach adopted, see Annex 1. In total, 1849 additional assessments are needed.

Table 9 provides an overview of the volumes indicated per stratum.

Methodology for assessing the nature of the company (SME or otherwise): additional searches (15)

Samples of applicants were compiled for each Member State on the basis of the sizes calculated above and randomised within the three quantiles.

We carried out web searches to assess the actual size of the applicants, using different entries to identify potential company homonyms, location information (country, street, city) and legal form indications.

Company entry on internet	Example
With/without legal form	Srl., Ltd., SA, …
With/without accompanying words	Research & development, service, consultancy,…
With/without country abbreviation	DE, BE, AT, …

In most cases, an appropriate website can be found. If an English version is available, this will be searched for size and/or dependency information (website sections: 'About us', 'Company', 'Investors', 'Profile', 'Financial data', 'Facts and Figures', 'Investors', 'Downloads' with annual reports). In the case of non-English websites, information is consulted in the native language (and, if necessary, translated by Google Translate). In the absence of unambiguous, relevant web pages, the case is classified as 'undecided'.

If insufficient information is found on the internet, a more exhaustive search is performed using available addresses in PATSTAT and/or group information and/or former names of the company found online. The information retrieved in this way is verified case by case in Amadeus 2012, filling the gaps resulting from reliance on exact matches only in the automated matching procedure. Overall, the combination of information from multiple sources yields an unambiguous assessment.

Information on size and/or dependency can be found for approximately 50 % of the names. In some cases, the organisation proves to be an individual, research institute or governmental agency ([16]). A number of cases (40 % +) require exhaustive searches, which can still result in their being classified as 'undecided'. The considerable number of undecided cases leaves three options for calculating the proportion of SMEs:

- ignore the undecided cases and calculate estimators based on identified cases only; this approach assumes no correlation between the size of a firm and online presence/absence; or
- build on the assumption that small companies will tend to invest less in web presence than large firms; the logical conclusion will be to classify all undecided cases as SMEs;
- alternatively, an extra 'undecided' category is extrapolated, creating a margin in which the proportion of SMEs to large companies may vary.

This final approach will be adopted in the results section.

Merging the results from automated firm-size classification and firm-size extrapolation

Table 8 presents the distribution of applicants and their corresponding patent volume per stratum across the four categories resulting from the matching, disambiguation and classification stages:

- 'large' companies;
- SMEs;
- non-matched corporate applicants; and
- matched corporate applicants lacking sufficient size/ownership information.

([15]) Depending on the language, 10 to 15 company names are identified per hour. Given the total of 1 849 corporate applicants to verify, we estimate that between 21 and 28 working days are sufficient to verify all EU-27 samples.

([16]) This negligible proportion of non-matched corporate applicants classed as non-corporate can be attributed to the error margin of the sector allocation algorithm.

Table 8: Results of stratification of applicant and patent population

Country	Quantile	Patent volume thresholds	Large		Small		Matched but unknown		Not matched	
			Applts.	Patents	Applts.	Patents	Applts.	Patents	Applts.	Patents
BE	1	> 702	5	6948	1	1210	-	-	-	-
BE	2	58 – 702	34	7417	1	59	4	420	7	1109
BE	3	< 58	359	2755	213	1100	925	2962	669	1800
BG	1	> 4	2	16	-	-	1	6	7	47
BG	2	1 – 4	22	25	4	6	16	20	55	70
BG	3	< 1	-	-	-	-	-	-	-	-
CZ	1	> 10	7	225	-	-	1	47	6	206
CZ	2	2 – 10	46	167	26	86	86	272	56	175
CZ	3	< 2	31	31	36	36	103	103	102	102
DK	1	> 402	9	8971	1	730	-	-	-	-
DK	2	21 – 402	68	6096	33	1426	22	851	36	1682
DK	3	< 21	280	1352	581	1897	1107	3132	1456	3337
DE	1	> 2544	19	125606	-	-	2	45006	2	6681
DE	2	147 – 2544	282	142005	11	2691	25	11730	62	24674
DE	3	< 147	4841	74112	2999	14268	8141	37831	13746	52746
EE	1	> 4	-	-	4	21	4	24	3	30
EE	2	2 – 4	3	10	9	20	10	26	12	28
EE	3	< 2	3	3	14	14	18	18	32	32
IE	1	> 81	7	1384	-	-	5	538	3	935
IE	2	12 – 81	26	948	8	173	52	1211	23	588
IE	3	< 12	80	226	56	156	678	1810	297	669
EL	1	> 11	1	14	1	11	2	33	4	177
EL	2	3 – 11	3	14	5	18	11	54	28	156
EL	3	< 3	5	8	8	12	23	32	118	147
ES	1	> 28	45	3586	9	359	12	703	18	1175
ES	2	4 – 28	174	1744	68	501	254	1806	293	2141
ES	3	< 4	297	471	233	349	1303	1972	1528	2209
FR	1	> 1602	12	55996	-	-	1	1698	1	1688
FR	2	102 – 1602	118	43355	5	848	16	3439	52	12138
FR	3	< 102	1514	19967	756	4701	3165	13454	5123	21519
IT	1	> 84	72	19587	4	606	3	637	19	4958
IT	2	9 – 84	471	11319	246	4346	327	5796	235	4894
IT	3	< 9	1212	3457	2695	5816	3944	8754	3876	6976
CY	1	> 16	1	19	-	-	5	165	6	127
CY	2	4 – 16	3	14	-	-	9	59	42	273
CY	3	< 4	1	1	6	9	37	56	135	209
LV	1	> 37	-	-	-	-	-	-	2	81
LV	2	4 – 37	1	6	1	6	-	-	7	96
LV	3	< 4	4	6	1	1	11	18	47	74
LT	1	> 3	-	-	-	-	1	3	2	8
LT	2	1 – 3	1	2	-	-	2	4	-	-
LT	3	< 1	-	-	2	2	2	2	6	6
LU	1	> 203	2	835	-	-	2	840	-	-
LU	2	16 – 203	8	352	-	-	4	304	26	1302
LU	3	< 16	54	190	40	113	149	466	364	990
HU	1	> 53	2	245	-	-	-	-	2	276
HU	2	3 – 53	3	52	2	6	26	141	65	442
HU	3	< 3	11	14	7	9	130	169	265	335
MT	1	> 28	1	29	-	-	1	112	-	-
MT	2	7 – 28	1	20	2	28	7	84	2	15
MT	3	< 7	3	10	4	8	34	72	27	48
NL	1	> 0	-	-	-	-	-	-	-	-
NL	2	1444 – 0	8	83241	-	-	-	-	1	3932
NL	3	< 1444	791	18586	485	2690	3436	15645	2170	7618
AT	1	> 120	18	6510	1	183	3	426	6	1271
AT	2	20 – 120	93	4520	18	614	37	1303	53	2194
AT	3	< 20	325	1401	269	839	868	2789	1351	3240
PL	1	> 25	5	205	2	108	-	-	3	92
PL	2	3 – 25	21	117	12	58	18	82	29	141
PL	3	< 3	47	59	38	47	95	120	131	150
PT	1	> 15	5	185	3	95	2	39	1	21
PT	2	2 – 15	26	98	24	100	41	130	56	173
PT	3	< 2	19	19	28	28	44	44	133	133
RO	1	> 3	-	-	-	-	3	16	6	22
RO	2	1 – 3	5	7	1	1	8	10	34	39
RO	3	< 1	-	-	-	-	-	-	-	-
SI	1	> 175	-	-	-	-	-	-	1	429
SI	2	8 – 175	5	81	-	-	7	268	9	127
SI	3	< 8	13	44	12	36	98	176	120	204
SK	1	> 8	3	42	1	12	2	41	-	-
SK	2	2 – 8	6	28	3	7	20	54	22	54
SK	3	< 2	7	7	6	6	28	28	26	26
FI	1	> 0	-	-	-	-	-	-	-	-
FI	2	225 – 0	13	31601	-	-	-	-	2	2358
FI	3	< 225	333	6386	304	1634	1074	5222	957	3820
SE	1	> 5098	1	6057	-	-	-	-	1	19205
SE	2	97 – 5098	55	25440	6	1074	4	1207	11	3531
SE	3	< 97	579	6406	491	4069	2316	8702	2762	9027
UK	1	> 225	52	40697	1	552	3	767	1	354
UK	2	16 – 225	436	22310	234	8021	317	11072	41	1736
UK	3	< 16	1895	7521	3474	10476	6981	19010	2876	4951

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Using the computation methodology outlined above, sample sizes are calculated for the latter two categories in the two left-hand panels of **Table 8**. Samples are then randomly drawn from the full, stratified populations of non-matched corporate applicants and matched corporate applicants for which available information is insufficient to determine size. The results of searches on corporate applicants in both categories across all Member States are reported in **Tables 9** and **10**.

Table 9: Sample sizes and results of sample checks per stratum
(matched applicants with unknown firm size)

Ctry	Quantile	Patent volume thresholds	Matched but size unknown										Total
			Large2		SME		Gov./research institute		Individual		Unknown		
			#	%	#	%	#	%	#	%	#	%	
BE	1	> 702											
BE	2	58 – 702	3	75	1	25		0		0		0	4
BE	3	< 58	3	17	11	61		0		0	4	22	18
BG	1	> 4		0	1	100		0		0		0	1
BG	2	1 – 4	1	20	4	80		0		0		0	5
BG	3	< 1											
CZ	1	> 10	1	100		0		0		0		0	1
CZ	2	2 – 10		0	5	100		0		0		0	5
CZ	3	< 2	1	20	4	80		0		0		0	5
DK	1	> 402											
DK	2	21 – 402	2	40	3	60		0		0		0	5
DK	3	< 21	2	10	16	76		0		0	3	14	21
DE	1	> 2 544	2	100		0		0		0		0	2
DE	2	147 – 2 544	4	80		0		0		0	1	20	5
DE	3	< 147	18	11	104	66		0		0	36	23	158
EE	1	> 4	1	25	3	75		0		0		0	4
EE	2	2 – 4	2	40	2	40		0		0	1	20	5
EE	3	< 2	1	20	4	80		0		0		0	5
IE	1	> 81	1	20	4	80		0		0		0	5
IE	2	12 – 81		0	4	80		0		0	1	20	5
IE	3	< 12		0	14	100		0		0		0	14
EL	1	> 11		0	2	100		0		0		0	2
EL	2	3 – 11	1	20	4	80		0		0		0	5
EL	3	< 3		0	5	100		0		0		0	5
ES	1	> 28	2	40	3	60		0		0		0	5
ES	2	4 – 28	2	40	3	60		0		0		0	5
ES	3	< 4		0	26	100		0		0		0	26
FR	1	> 1 602	1	100		0		0		0		0	1
FR	2	102 – 1 602	3	60	2	40		0		0		0	5
FR	3	< 102	6	10	53	88		0		0	1	2	60
IT	1	> 84	3	100		0		0		0		0	3
IT	2	9 – 84	1	14	6	86		0		0		0	7
IT	3	< 9	4	5	68	93		0		0	1	1	73
CY	1	> 16	3	60	2	40		0		0		0	5
CY	2	4 – 16		0	5	100		0		0		0	5
CY	3	< 4		0	5	100		0		0		0	5
LV	1	> 37											
LV	2	4 – 37											
LV	3	< 4		0	5	100		0		0		0	5
LT	1	> 3		0	1	100		0		0		0	1
LT	2	1 – 3	1	50	1	50		0		0		0	2
LT	3	< 1		0	2	100		0		0		0	2
LU	1	> 203		0	2	100		0		0		0	2
LU	2	16 – 203		0	4	100		0		0		0	4
LU	3	< 16		0	5	100		0		0		0	5
HU	1	> 53											
HU	2	3 – 53	1	20	4	80		0		0		0	5
HU	3	< 3	1	20	3	60		0	1	20		0	5
MT	1	> 28		0	1	100		0		0		0	1
MT	2	7-28	1	20	4	80		0		0		0	5
MT	3	< 7	1	20	4	80		0		0		0	5
NL	1	> 0											
NL	2	1 444 – 0											
NL	3	< 1 444	7	10	56	84		0		0	4	6	67
AT	1	> 120	2	67	1	33		0		0		0	3
AT	2	20 – 120	4	80	1	20		0		0		0	5
AT	3	< 20	3	17	15	83		0		0		0	18
PL	1	> 25											
PL	2	3 – 25	1	20	3	60		0		0	1	20	5
PL	3	< 3	2	40	2	40		0		0	1	20	5
PT	1	> 15		0	1	50		0		0	1	50	2
PT	2	2 – 15	1	20	4	80		0		0		0	5
PT	3	< 2		0	4	80		0		0	1	20	5
RO	1	> 3	1	33	1	33		0		0	1	33	3
RO	2	1 – 3		0	5	100		0		0		0	5
RO	3	< 1											
SI	1	> 175											
SI	2	8 – 175	3	60	2	40		0		0		0	5
SI	3	< 8		0	5	100		0		0		0	5
SK	1	> 8		0	2	100		0		0		0	2
SK	2	2 – 8	1	20	4	80		0		0		0	5
SK	3	< 2	1	20	4	80		0		0		0	5
FI	1	> 0											
FI	2	225 – 0											
FI	3	< 225	5	23	15	68		0		0	2	9	22
SE	1	> 5 098											
SE	2	97 – 5 098	4	100		0		0		0		0	4
SE	3	< 97	6	13	38	81		0	2	4	1	2	47
UK	1	> 225	1	33	1	33	1	33		0		0	3
UK	2	16 – 225	2	29	5	71		0		0		0	7
UK	3	< 16	20	15	102	76	1	1		0	11	8	134

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Table 10: Sample sizes and results of sample checks per stratum (non-matched applicants)

Ctry	Quantile	Patent volume thresholds	Not matched										Total
			Large		SME		Gov./research institute owned		Individual		Unknown		
			#	%	#	%	#	%	#	%	#	%	
BE	1	> 702											
BE	2	58 – 702	5	100		0		0		0		0	5
BE	3	< 58	10	77	3	23		0		0		0	13
BG	1	> 4	1	20	4	80		0		0		0	5
BG	2	1 – 4		0	1	20	1	20	2	40	1	20	5
BG	3	< 1											
CZ	1	> 10	4	80		0		0		0	1	20	5
CZ	2	2 – 10	5	100		0		0		0		0	5
CZ	3	< 2	2	40	3	60		0		0		0	5
DK	1	> 402											
DK	2	21 – 402	5	100		0		0		0		0	5
DK	3	< 21	7	25	9	32		0	3	11	9	32	28
DE	1	> 2 544	2	100		0		0		0		0	2
DE	2	147 – 2 544	5	100		0		0		0		0	5
DE	3	< 147	116	43	68	25	3	1	5	2	75	28	267
EE	1	> 4		0	3	100		0		0		0	3
EE	2	2 – 4	1	20	4	80		0		0		0	5
EE	3	< 2	1	20	4	80		0		0		0	5
IE	1	> 81	3	100		0		0		0		0	3
IE	2	12 – 81	4	80		0	1	20		0		0	5
IE	3	< 12	2	33	2	33		0		0	2	33	6
EL	1	> 11	4	100		0		0		0		0	4
EL	2	3 – 11		0	3	60		0	2	40		0	5
EL	3	< 3	2	40		0		0		0	3	60	5
ES	1	> 28	5	100		0		0		0		0	5
ES	2	4 – 28	4	67	1	17		0		0	1	17	6
ES	3	< 4	12	39	11	35	1	3	3	10	4	13	31
FR	1	> 1 602	1	100		0		0		0		0	1
FR	2	102 – 1 602	5	100		0		0		0		0	5
FR	3	< 102	44	45	28	29	4	4	2	2	19	20	97
IT	1	> 84	5	100		0		0		0		0	5
IT	2	9 – 84	4	80	1	20		0		0		0	5
IT	3	< 9	17	24	38	53		0	1	1	16	22	72
CY	1	> 16	3	60	2	40		0		0		0	5
CY	2	4 – 16	1	20	4	80		0		0		0	5
CY	3	< 4		0	3	60		0		0	2	40	5
LV	1	> 37	1	50		0		0	1	50		0	2
LV	2	4 – 37		0	1	20		0	4	80		0	5
LV	3	< 4		0	3	60		0	1	20	1	20	5
LT	1	> 3	1	50		0	1	50		0		0	2
LT	2	1 – 3											
LT	3	< 1	2	40	1	20	1	20	1	20		0	5
LU	1	> 203											
LU	2	16 – 203	4	80	1	20		0		0		0	5
LU	3	< 16	2	25	1	13		0		0	5	63	8
HU	1	> 53	2	100		0		0		0		0	2
HU	2	3 – 53	2	40	3	60		0		0		0	5
HU	3	< 3	1	17	1	17		0	3	50	1	17	6
MT	1	> 28											
MT	2	7 – 28		0	2	100		0		0		0	2
MT	3	< 7	1	20	3	60		0		0	1	20	5
NL	1	> 0											
NL	2	1 444 – 0	1	100		0		0		0		0	1
NL	3	< 1 444	15	35	19	44		0	2	5	7	16	43
AT	1	> 120	5	100		0		0		0		0	5
AT	2	20 – 120	5	100		0		0		0		0	5
AT	3	< 20	13	48	7	26	1	4	2	7	4	15	27
PL	1	> 25	2	67	1	33		0		0		0	3
PL	2	3 – 25	3	60		0	2	40		0		0	5
PL	3	< 3	1	20		0	2	40	2	40		0	5
PT	1	> 15		0	1	100		0		0		0	1
PT	2	2 – 15	3	60	1	20		0		0	1	20	5
PT	3	< 2		0	2	40		0	2	40	1	20	5
RO	1	> 3	2	40	3	60		0		0		0	5
RO	2	1 – 3	3	60	2	40		0		0		0	5
RO	3	< 1											
SI	1	> 175	1	100		0		0		0		0	1
SI	2	8 – 175	4	80	1	20		0		0		0	5
SI	3	< 8	1	20	3	60		0		0	1	20	5
SK	1	> 8											
SK	2	2 – 8	3	60	1	20		0		0	1	20	5
SK	3	< 2	1	20	2	40	1	20		0	1	20	5
FI	1	> 0											
FI	2	225 – 0	2	100		0		0		0		0	2
FI	3	< 225	6	30	8	40		0		0	6	30	20
SE	1	> 5 098	1	100		0		0		0		0	1
SE	2	97 – 5 098	5	100		0		0		0		0	5
SE	3	< 97	18	33	22	40		0	4	7	11	20	55
UK	1	> 225	1	100		0		0		0		0	1
UK	2	16 – 225	1	20	2	40		0		0	2	40	5
UK	3	< 16	13	24	26	47	1	2	3	5	12	22	55

Source: PATSTAT autumn 2011 edition, Amadeus 2012, Internet searches based on applicant name.

While the web search approach to classification into large companies and SMEs seems to have potential for the top two quantiles, this is not so clear for the third (comprising applicants with the fewest patent applications). A significant amount of firm-size information remains unknown — in particular, the matched information for automated classification.

To extrapolate the findings from the samples, the percentages of large companies, SMEs, non-corporate applicants and unknown-size companies are calculated per stratum. Next, these percentages (see **Tables 9** and **10**) are multiplied by the corresponding number of applicants per stratum in the population (see **Table 8**). This produces per-population stratum estimates of the numbers of:

- identifiable SMEs;
- large companies;
- non-corporate applicants; and
- companies of which the size remains unidentified as a result of sole reliance on web searches.

Finally, the numbers of 'large' companies and SMEs identified using the automated matching, disambiguation and classification procedures are added to the estimates per stratum, and stratum counts are aggregated per 'size type' to obtain an enhanced picture of the proportion of corporate applicants that are SMEs.

The results combining extrapolated numbers with the automated outcomes are presented, according to proportion, in Annex 2 (**Table 21**). Of the EU applicants that filed patents from 1999 onwards, 34% appear to be large companies and 54% SMEs. 13% are of unknown size. The proportions of SMEs among companies in the UK, Italy, Ireland, Slovenia, Latvia, Estonia, Malta and Cyprus filing for patent protection appear to be higher than the EU altogether ([17]). Caution is called for, though, when interpreting these underlying country proportions, as sample sizes are optimised to obtain 95% precision for the overall EU SME proportion.

To establish the proportion of SME patents in the total population of patents filed by EU companies, the sample proportions (see **Tables 9** and **10**) are multiplied by population patent volumes per stratum (see **Table 8**) to produce per-stratum estimates of patent volume by firm-size type. As in the approach used to identify size-type proportion at applicant level, patent volumes attributed to applicants identified as SMEs or large companies in the automated matching stage are then added to the patent volumes estimated across strata and aggregated per size-type and per country. Results combining extrapolated numbers with the matching outcome in terms of patent volumes are presented in **Table 11**.

([17]) For these member states the share of SMEs among the patenting firms lies above the combined overall EU share of SMEs and firms with unknown size.

Table 11: Extrapolation to population: applications
(%)

Country	Large	Unknown — non-matched	Unknown — matched	SME
EU-27	**78.9**	**2.3**	**1.2**	**17.6**
BE	79.2	0.0	2.6	18.2
BG	36.8	9.5	0.0	53.8
CZ	60.1	2.8	0.0	37.1
DK	67.2	3.7	1.5	27.6
DE	84.9	2.8	2.0	10.3
EE	19.9	0.0	2.3	77.8
IE	50.4	2.6	2.8	44.1
EL	46.1	14.4	0.0	39.6
ES	61.3	3.8	0.0	34.8
FR	83.4	2.4	0.1	14.1
IT	60.8	2.0	0.2	37.1
CY	28.3	9.0	0.0	62.7
LV	33.7	9.5	0.0	56.8
LT	50.5	0.0	0.0	49.5
LU	49.4	11.5	0.0	39.1
HU	59.3	3.8	0.0	37.0
MT	23.4	2.3	0.0	74.3
NL	83.8	0.9	0.7	14.6
AT	77.2	1.9	0.0	20.9
PL	62.0	0.0	4.0	34.0
PT	42.7	6.0	2.8	48.5
RO	46.9	0.0	5.6	47.5
SI	62.8	3.0	0.0	34.2
SK	43.7	5.3	0.0	51.0
FI	83.6	2.2	0.9	13.2
SE	78.8	2.2	0.2	18.9
UK	62.1	1.4	1.2	35.3

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Table 11 shows that 3.5% of patent volume can be assigned to corporate applicants of unknown size. 17.6% of the applications originate from innovative activities in SMEs, whereas 78.9% are filed by 'large' companies. More established knowledge economies such as Germany, France, The Netherlands, Sweden, Belgium, Austria, Finland – that are not by coincidence known to host headquarters of some of the bigger multinationals – tend to show lower proportions of patents filed by SMEs than in the EU overall. However exceptions to this observed tendency can be reported as well: equivalent to some of the more peripheral and more recent Member States, equally advanced economies like the United Kingdom, Denmark, Italy and Spain show SME contributions above the overall EU-level. Again, note that the country proportions presented here should be merely regarded as indicative as the adopted sampling strategy was designed to infer the EU overall proportion.

3. Further analysis

The data presented in the previous chapter suggest the need for a thorough assessment of SMEs' contribution to patenting activity per country. While such country-level analyses are relevant, they do not establish whether and how SMEs' contribution varies across technologies. In this section, we identify where and to what extent SMEs are more active.

Patents are allocated to technology fields on the basis of the International Patent Classification code(s) assigned to applications at the time of their examination. A translation of the Fraunhofer classification scheme developed by Schmoch *et al.* (2008) was used to classify patent applications in 35 fields.

In the following analyses, we provide a number of descriptive statistics revealing the relative contribution of SME applicants per field of technology. Next, we compute technological specialisation profiles per Member State and introduce multivariate analyses to assess whether countries' technological specialisation in a given field is driven by large, established firms or by smaller business entities.

In this analysis, we focus on EPO and PCT patent applications from, and USPTO patents granted to, companies for which a match was found in Amadeus. We consider the time frame 2005 to 2011.

SMEs' contribution per technology

For each technology *t,* the overall proportion of patent applications filed (by companies with identified firm size) that came from SMEs is computed as follows:

$$SME\ contribution_t = \frac{\#APPLNS_{SME,t}}{\#APPLNS_{SME,t} + \#APPLNS_{LARGE,t}}$$

The overall SME contribution per technology field is reported in the left-hand panel of **Table 12**. The highest contribution of SMEs, a 'market share' of more than 20 %, is in biotechnology-related fields (including Biotechnology [15], Pharmaceuticals [16] and the Analysis of Biological Materials [11]). For a number of ICT domains (Telecommunications [3], Digital Communication [4]) and Macromolecular Chemistry [17], the SME contribution is modest (less than 4 %).

Table 12: SMEs' contribution to corporate technology development, EU-27

FhG35 area	FhG35 code	FhG35 field of technology	SME patents among patents filed by companies with known firm size (%)	*Rank*
Electrical engineering	1	Electrical machinery, apparatus, energy	6.0	*24*
	2	Audio-visual technology	5.2	*28*
	3	Telecommunications	3.4	*34*
	4	Digital communication	3.0	*35*
	5	Basic communication processes	3.6	*32*
	6	Computer technology	6.0	*25*
	7	IT methods for management	13.0	*6*
	8	Semiconductors	6.9	*21*
Instruments	9	Optics	7.1	*19*
	10	Measurement	7.0	*20*
	11	Analysis of biological materials	23.3	*1*
	12	Control	6.7	*22*
	13	Medical technology	14.3	*4*
Chemistry	14	Organic fine chemistry	4.5	*29*
	15	Biotechnology	20.7	*2*
	16	Pharmaceuticals	17.5	*3*
	17	Macromolecular chemistry. polymers	3.5	*33*
	18	Food chemistry	12.8	*7*
	19	Basic materials chemistry	5.3	*27*
	20	Materials. metallurgy	6.2	*23*
	21	Surface technology. coating	8.2	*18*
	22	Micro-structure and nano-technology	10.7	*10*
	23	Chemical engineering	10.1	*11*
	24	Environmental technology	8.8	*17*
Mechanical engineering	25	Handling	11.4	*9*
	26	Machine tools	9.3	*15*
	27	Engines, pumps, turbines	3.7	*31*
	28	Textile and paper machines	9.5	*14*
	29	Other special machines	10.0	*12*
	30	Thermal processes and apparatus	9.6	*13*
	31	Mechanical elements	5.7	*26*
	32	Transport	4.0	*30*
Other fields	33	Furniture, games	14.1	*5*
	34	Other consumer goods	9.2	*16*
	35	Civil engineering	12.8	*8*

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Table 13 provides a more fine grained image of SMEs' relative contribution to certain technology fields across Member States ([18]). A 'relative SME contribution' indicator is introduced whereby the proportion of patent applications filed by SMEs from country *c* in technology *t* is divided by the proportion of patent applications filed by SMEs in country *c* overall:

$$Relative\ SME\ Contribution_{c,t} = \frac{\frac{\#APPLNS_{SME,c,t}}{\#APPLNS_{SME,t,c} + \#APPLNS_{LARGE,c,t}}}{\frac{\#APPLNS_{SME,c,t}}{\#APPLNS_{SME,c,t} + \#APPLNS_{LARGE,c,t}}}$$

Values above 1 reflect a relatively stronger contribution from SMEs, while those below 1 (with a minimum of zero) signal a less pronounced contribution. As this scale is asymmetrical (ranging from 0 to +∞), the indicator obtained was rescaled to values ranging from -1 to 1. Values near 1 suggest a relatively strong contribution from SMEs, while values close to -1 signal a marginal contribution ([19]).

([18]) Bulgaria and Latvia were discarded, since too few patenting SMEs were found among the companies matched to corporate applicants.

([19]) Relative SME contribution rates per country c and technology *t* are rescaled as follows:

$$Rescaled\ relative\ SME\ contribution_{c,t} = \frac{Relative\ SME\ contribution_{c,t} - 1}{Relative\ SME\ contribution_{c,t} + 1}$$

Table 13: Relative contribution of SMEs to corporate technology development — per country & field

(Higher values indicate higher levels of contribution)

FhG35	Field of technology	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	IT	CY	LV	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK
1	Electrical machinery, apparatus, energy	-0.09	-1.00	-0.32	-0.28	-0.20	0.07	0.68	0.42	-0.04	-0.24	-0.11	-1.00	-1.00	-0.31	0.29	0.09	0.00	-0.22	-0.23	0.07	-1.00	0.24	0.22	-0.05	-0.02	0.10
2	Audio-visual technology	0.07	0.15	0.38	-0.31	-0.01	0.07	-0.61	-1.00	-0.28	-0.12	0.05	0.06	-1.00	0.37	-1.00	-1.00	-0.76	-0.01	-1.00	0.27	-1.00	-1.00	0.22	-0.22	-0.23	-0.02
3	Telecommunications	0.62	0.15	0.38	-0.04	-0.23	0.07	0.09	0.42	-0.22	-0.57	-0.41	-0.09	-1.00	-0.01	-1.00	0.14	-0.62	-0.10	0.42	-0.03	0.32	0.84	0.22	-0.77	-0.77	-0.11
4	Digital communication	0.76	0.15	-0.06	0.07	-0.24	0.07	0.07	-1.00	0.03	-0.48	-0.63	-1.00	-1.00	-1.00	0.76	0.14	-0.72	0.10	-1.00	0.35	0.32	-1.00	0.22	-0.77	-0.85	-0.20
5	Basic communication processes	0.72	-1.00	-1.00	-0.64	-0.27	-1.00	0.33	-1.00	0.46	-0.21	-0.21	-1.00	-1.00	-1.00	-1.00	-1.00	-0.88	-1.00	-1.00	-1.00	-1.00	0.84	-1.00	-0.84	-0.19	-0.14
6	Computer technology	0.26	0.15	-0.39	0.14	-0.15	0.07	-0.60	0.42	-0.08	0.05	-0.16	-0.49	-1.00	-0.06	-1.00	0.08	-0.63	0.17	0.22	0.07	-1.00	0.86	0.22	-0.24	0.03	-0.12
7	IT methods for management	0.72	-1.00	0.52	0.28	0.01	0.07	-0.54	0.42	0.39	-0.19	0.22	-1.00	-1.00	-0.18	-1.00	0.14	0.25	-0.27	-1.00	-0.19	0.32	0.88	0.22	0.12	-0.16	0.24
8	Semiconductors	-0.17	-1.00	-1.00	0.16	-0.01	-1.00	-1.00	0.42	-0.43	0.27	-0.02	0.09	-1.00	-0.68	-1.00	-1.00	-0.37	-0.30	0.42	0.35	-1.00	0.88	-1.00	0.52	0.05	0.28
9	Optics	-0.52	0.15	0.38	0.07	-0.05	0.07	-0.02	-1.00	-1.00	0.26	-0.29	0.23	-1.00	0.44	-1.00	0.14	-0.42	0.01	-0.34	0.35	-1.00	0.63	-1.00	0.40	0.37	0.02
10	Measurement	0.36	-1.00	0.09	0.21	0.00	-0.07	0.15	-1.00	0.03	0.01	0.00	0.03	-1.00	-0.86	-1.00	-0.81	-0.47	0.05	-0.13	0.01	-1.00	0.41	-0.12	0.17	0.18	-0.08
11	Analysis of biological materials	0.53	-1.00	-1.00	0.31	0.56	0.07	-0.11	0.42	0.35	0.52	0.31	0.23	-1.00	0.22	0.76	0.14	0.51	0.60	-1.00	0.35	-1.00	-1.00	-1.00	0.69	0.72	0.26
12	Control	0.14	0.15	0.03	0.13	-0.30	0.07	-0.09	0.42	0.08	-0.23	0.08	-1.00	-1.00	-1.00	0.76	-1.00	0.07	-0.52	-0.09	-0.03	-1.00	-1.00	0.22	0.28	-0.17	0.05
13	Medical technology	0.66	-0.38	0.13	-0.39	0.36	-0.04	0.10	-1.00	0.24	0.65	0.18	0.03	-1.00	0.47	0.71	0.02	-0.09	0.33	0.10	0.35	0.32	-1.00	0.22	0.55	0.19	0.17
14	Organic fine chemistry	-0.50	0.15	-0.18	-0.01	-0.22	-0.02	-0.10	-0.16	-0.26	-0.33	-0.39	0.01	-1.00	0.43	-0.47	-0.27	-0.37	-0.14	-1.00	-0.40	-1.00	-1.00	-1.00	0.23	0.23	-0.39
15	Biotechnology	0.27	0.15	-0.23	-0.16	0.57	-0.15	-1.00	0.42	0.48	0.73	0.20	0.23	-1.00	0.17	0.38	0.14	0.62	0.49	0.49	0.13	-1.00	0.28	-1.00	0.69	0.62	0.29
16	Pharmaceuticals	0.01	-1.00	0.36	0.31	0.41	-0.04	-0.08	-1.00	-0.22	0.57	-0.20	0.08	-1.00	0.46	-0.73	-0.58	0.50	0.31	-0.27	-0.05	-0.01	-1.00	-0.63	0.78	0.43	0.11
17	Macromolecular chemistry, polymers	-0.61	-1.00	0.51	-0.50	-0.27	0.07	-0.85	0.42	-0.08	-0.59	-0.68	0.23	-1.00	0.25	0.76	0.14	-0.28	-0.14	-1.00	-0.10	-1.00	-1.00	-1.00	-0.68	0.45	-0.07
18	Food chemistry	0.27	-1.00	0.63	-0.25	0.56	-0.13	0.31	-1.00	0.47	0.55	0.16	-1.00	-1.00	-1.00	0.42	-1.00	0.31	-0.16	-0.01	0.35	-1.00	-1.00	0.22	0.33	0.78	-0.26
19	Basic materials chemistry	-0.37	-1.00	0.48	-0.22	-0.15	-0.10	-0.28	0.42	0.21	-0.16	-0.32	0.15	-1.00	-0.05	0.10	0.14	-0.36	-0.14	-0.06	-0.26	-1.00	-1.00	-1.00	-0.28	0.52	-0.24
20	Materials, metallurgy	-0.12	0.15	0.28	0.09	-0.08	-1.00	-0.66	-1.00	0.01	-0.04	-0.18	-1.00	-1.00	-0.81	0.57	0.14	-0.63	-0.44	0.13	-0.59	0.32	-1.00	-0.31	0.55	0.00	-0.21
21	Surface technology, coating	-0.42	-1.00	0.48	-0.08	-0.03	0.07	-1.00	0.11	0.00	0.01	-0.02	0.08	-1.00	-0.05	-1.00	0.13	0.12	-0.18	0.37	-0.10	-0.12	-1.00	-1.00	0.67	-0.03	-0.04
22	Micro-structure and nano-technology	0.58	-1.00	-1.00	0.13	-0.43	-1.00	-1.00	-1.00	0.70	0.08	-0.27	-1.00	-1.00	0.17	-1.00	-1.00	0.11	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	0.62	0.82	0.22
23	Chemical engineering	0.00	-1.00	0.42	0.11	0.10	-0.02	0.26	0.28	0.06	-0.12	0.18	0.23	-1.00	-0.08	0.61	-0.27	0.18	0.31	0.26	-0.25	-0.21	-1.00	-1.00	0.35	0.43	-0.08
24	Environmental technology	-0.04	-1.00	0.15	0.19	-0.01	-1.00	0.67	0.42	-0.15	-0.20	0.15	-1.00	-1.00	-0.86	0.76	0.14	0.12	0.07	0.45	0.24	0.32	-1.00	0.22	0.52	0.26	-0.10
25	Handling	-0.04	-1.00	0.22	0.08	0.11	-0.27	-0.66	0.42	0.06	0.03	0.09	0.09	-1.00	-0.03	0.76	-0.06	0.50	0.01	0.22	0.26	0.32	0.85	-1.00	-0.19	0.12	0.02
26	Machine tools	0.13	-1.00	-0.06	0.27	0.11	-1.00	-0.49	-0.77	-0.17	0.07	0.19	-0.06	-1.00	-0.62	-1.00	0.12	0.64	-0.38	-0.34	-0.04	-1.00	-1.00	-0.31	0.55	0.12	-0.09
27	Engines, pumps, turbines	-0.11	-1.00	-1.00	-0.77	-0.56	-1.00	0.74	0.11	-0.48	-0.75	-0.22	-1.00	-1.00	-0.67	-1.00	-1.00	0.51	-0.22	-1.00	0.01	-1.00	0.68	-1.00	0.28	-0.13	-0.19
28	Textile and paper machines	-0.67	-1.00	-0.83	0.32	0.25	0.07	-0.59	-1.00	0.53	-0.26	0.09	0.16	-1.00	-0.40	-1.00	0.08	0.10	-0.10	-1.00	-0.19	-1.00	-1.00	-1.00	-0.18	0.22	0.06
29	Other special machines	-0.28	-1.00	0.09	0.04	0.08	0.07	0.09	0.23	-0.12	0.20	0.03	0.23	-1.00	-0.31	-1.00	-0.02	0.49	0.08	0.19	0.06	-1.00	0.77	0.22	0.40	-0.08	0.09
30	Thermal processes and apparatus	-0.59	-1.00	0.19	-0.24	0.06	0.07	0.71	-1.00	-0.06	0.02	0.00	-0.11	-1.00	-0.58	0.76	0.14	0.53	0.15	-0.33	-0.26	-1.00	0.17	0.22	0.32	-0.27	0.17
31	Mechanical elements	0.00	0.15	0.08	-0.17	-0.32	-1.00	0.02	-1.00	0.03	-0.53	0.06	-1.00	-1.00	-0.90	-1.00	0.09	0.60	-0.23	0.23	-0.36	-1.00	0.65	0.19	0.03	-0.32	-0.07
32	Transport	-0.30	-1.00	-0.85	0.19	-0.51	0.07	0.35	0.42	-0.34	-0.69	-0.04	-0.06	-1.00	-0.91	-1.00	-1.00	0.45	0.13	0.35	0.19	0.32	0.51	-0.73	0.50	-0.52	-0.06
33	Furniture, games	-0.15	0.15	-0.01	0.31	0.10	0.07	-0.34	-1.00	0.33	0.16	0.11	-1.00	-1.00	-1.00	-1.00	0.14	0.48	-0.44	-0.41	0.26	-1.00	0.34	-1.00	0.52	0.38	0.17
34	Other consumer goods	-0.72	-1.00	0.12	-0.05	-0.02	-1.00	0.47	0.42	0.18	0.02	0.06	-0.41	-1.00	-0.80	-1.00	0.14	0.07	-0.42	-0.69	0.26	0.32	-1.00	0.22	0.43	0.41	0.01
35	Civil engineering	0.11	0.15	-0.03	-0.20	0.32	0.07	-0.53	0.42	0.23	0.08	0.20	0.05	-1.00	-0.72	-1.00	0.14	0.24	-0.17	-0.62	-0.51	0.32	-1.00	-1.00	0.48	-0.17	0.11

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Overall, these findings confirm the picture for the EU as a whole (see Table 13), i.e. relatively strong SME participation in biotechnology-related fields and the opposite in (a number of) ICT domains.

The relative technological advantage of nations: do SMEs contribute?

Next, we examine whether and to what extent the overall technological advantage of countries in relative terms coincides with the comparative position of SMEs. The specialisation of national innovation systems (NISs) in certain technologies is measured by dividing the proportion of applications pertaining to technology *t* for country *c* by the proportion pertaining to technology *t* for the EU overall:

$$Relative\ Technological\ Advantage_{c,t} = \frac{\#APPLNS_{c,t}}{\#APPLNS_c} \Big/ \frac{\#APPLNS_t}{\#APPLNS}$$

As with the relative SME contribution rates (see above), the relative technological advantages (RTAs) are rescaled to obtain values between 1 and -1. The closer the RTA value is to 1, the more a country is specialised in technology *t*.

Table 14 presents RTAs for Member States' NISs on the basis of patent applications filed by applicants from all active sectors per country. **Table 15** shows RTAs based on counts of patent applications filed only by identified SMEs and individuals. **Table 16** reports RTAs obtained by taking into account only patent applications that originate from companies identified as large. These are assumed to capture the relatively more routine innovative activities undertaken by incumbent R&D facilities, whereas the RTAs based on combined counts of patents filed by SMEs and individuals are supposed to capture innovation resulting from greater entrepreneurial initiative ([20]).

([20]) The addition to the class of entrepreneurial RTAs of patents filed by individuals resulted in the removal of PCT applications. Due to the double allocation of inventor names in the PCT system (to the class of inventors and the class of applicants), the inclusion of PCT patent counts introduces bias to the computation of entrepreneurial RTAs.

Table 14: Rescaled RTAs for national innovation systems as a whole — per country & field

(Higher values indicate higher levels of specialisation).

FhG35	Field of technology	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	IT	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK
1	Electrical machinery, apparatus, energy	-0.32	-0.36	0.07	-0.16	0.08	-0.10	-0.28	-0.21	-0.12	0.00	-0.04	-0.42	-0.73	-1.00	-0.12	-0.31	-0.42	0.06	0.11	-0.20	-0.60	-0.44	-0.23	0.09	-0.20	-0.39	-0.11
2	Audio-visual technology	-0.24	0.00	-0.62	0.20	-0.21	-0.67	0.26	-0.53	-0.32	-0.03	-0.43	0.01	0.46	0.01	0.06	-0.37	-0.30	0.51	-0.15	-0.38	-0.23	-0.42	-0.32	-0.32	0.10	-0.04	-0.03
3	Telecommunications	-0.51	-0.06	-0.57	-0.58	-0.33	-0.63	0.03	-0.61	-0.16	0.15	-0.41	0.11	-0.44	-1.00	-0.33	-0.95	-0.68	0.06	-0.53	-0.07	-0.62	-0.35	-0.37	0.00	0.62	0.55	0.03
4	Digital communication	-0.55	-1.00	-0.59	-0.71	-0.32	0.13	0.14	-0.74	-0.35	0.24	-0.43	-0.74	-0.22	-0.01	-0.46	-0.77	-0.53	0.03	-0.64	-0.73	-0.92	0.08	-0.58	-0.34	0.62	0.50	-0.02
5	Basic communication processes	-0.19	-1.00	0.25	-0.41	-0.08	-0.36	-0.21	-0.34	-0.59	0.10	-0.17	-0.45	0.60	-1.00	-0.40	-0.80	-1.00	0.30	-0.19	-0.53	-0.51	0.33	-0.20	-0.42	0.20	0.11	-0.12
6	Computer technology	-0.22	-0.17	-0.41	-0.36	-0.11	-0.29	0.42	-0.18	-0.39	0.06	-0.32	0.18	-0.23	-0.35	-0.10	-0.20	-0.42	0.21	-0.34	-0.35	-0.21	0.15	-0.67	-0.16	0.36	0.13	0.11
7	IT methods for management	-0.46	0.06	0.09	-0.35	-0.13	0.45	0.75	0.16	0.09	-0.02	-0.32	0.70	0.20	0.80	0.31	-0.09	-0.35	-0.14	-0.12	0.40	0.38	0.57	0.02	0.77	0.32	0.09	0.31
8	Semiconductors	0.21	0.05	-0.77	-0.57	0.04	0.03	-0.36	-0.50	-0.47	0.04	-0.16	0.09	-0.46	-1.00	-0.21	-0.24	0.03	0.29	0.18	-0.37	-0.38	-1.00	-0.66	-0.45	-0.41	-0.58	-0.25
9	Optics	0.12	-0.24	-0.32	-0.26	-0.08	-0.62	-0.13	-0.19	-0.38	-0.06	-0.36	-0.14	-0.11	0.62	0.39	0.03	-0.77	0.46	-0.26	-0.36	-0.64	-1.00	-0.20	-0.66	-0.22	-0.23	-0.04
10	Measurement	-0.21	-0.45	-0.13	-0.18	0.04	0.04	-0.27	-0.03	-0.27	-0.02	-0.15	-0.28	-0.49	0.35	-0.12	-0.13	0.02	0.10	-0.11	-0.03	-0.16	0.31	-0.17	-0.24	-0.12	-0.10	0.07
11	Analysis of biological materials	0.26	0.34	0.17	0.30	-0.14	0.60	0.15	0.50	0.19	0.07	-0.24	0.18	-1.00	0.02	-0.50	0.10	-0.64	-0.08	0.05	-0.16	0.24	0.25	-0.28	0.64	-0.10	0.05	0.27
12	Control	-0.38	-0.20	0.18	-0.32	0.05	0.51	0.06	-0.07	0.04	0.01	0.03	0.30	-0.39	-1.00	-0.24	-0.04	-0.40	-0.16	0.10	-0.32	0.11	0.24	-0.34	-0.38	-0.14	-0.02	0.04
13	Medical technology	-0.20	-0.12	-0.06	0.33	-0.05	0.18	0.48	0.01	-0.05	-0.17	0.00	0.12	-0.09	0.48	-0.07	0.11	-0.50	0.15	-0.01	0.18	0.02	0.28	-0.20	-0.65	-0.41	0.15	0.13
14	Organic fine chemistry	0.11	-0.09	0.34	-0.04	0.05	0.23	-0.27	-0.19	0.14	0.17	-0.18	-0.32	0.73	-0.08	-0.53	0.48	-0.11	-0.17	-0.46	0.30	0.27	-1.00	0.58	-0.08	-0.54	-0.52	0.00
15	Biotechnology	0.36	-0.03	0.17	0.59	-0.11	0.75	0.09	0.19	0.26	0.00	-0.30	0.25	0.11	0.70	-0.79	0.18	-1.00	0.00	0.03	0.27	0.43	-1.00	0.26	0.36	-0.25	-0.23	0.13
16	Pharmaceuticals	0.33	0.47	0.28	0.41	-0.20	0.25	0.35	0.21	0.42	0.04	0.02	-0.07	0.35	0.42	0.04	0.58	0.06	-0.25	0.03	0.23	0.39	0.16	0.73	0.07	-0.57	0.17	0.22
17	Macromolecular chemistry, polymers	0.41	-0.27	-0.43	-0.49	0.10	-0.64	-0.13	0.05	-0.47	-0.10	-0.03	-0.38	-1.00	0.08	-0.66	-0.22	-0.01	0.08	0.00	-0.32	-0.13	-1.00	-0.81	-0.26	0.11	-0.70	-0.27
18	Food chemistry	0.34	0.41	-0.38	0.60	-0.22	0.20	0.00	0.30	0.27	-0.21	-0.07	-0.05	0.32	0.33	-0.21	0.14	-1.00	0.48	-0.42	-0.07	0.16	-1.00	-0.11	-0.21	-0.41	-0.62	-0.01
19	Basic materials chemistry	0.18	0.14	0.03	0.00	0.10	-0.18	-0.26	-0.10	-0.10	-0.11	-0.24	-0.13	-0.25	-1.00	-0.40	-0.05	-0.02	0.07	-0.18	-0.04	-0.19	-0.46	-0.55	0.21	-0.37	-0.57	0.07
20	Materials, metallurgy	0.19	0.18	0.12	-0.15	0.03	0.38	-0.12	-0.07	0.02	0.09	-0.10	0.15	-0.02	-0.15	0.54	0.03	-0.13	-0.31	0.29	0.32	-0.09	0.08	-0.50	0.55	-0.10	-0.10	-0.11
21	Surface technology, coating	0.21	0.24	-0.26	-0.22	0.06	-0.09	-0.12	0.34	-0.24	0.02	-0.08	0.21	-0.47	-1.00	0.24	-0.21	0.54	-0.15	0.00	-0.05	-0.16	0.28	-0.33	-0.46	-0.03	-0.11	-0.06
22	Micro-structure and nano-technology	0.10	-1.00	-0.25	-0.36	-0.12	-1.00	-0.04	0.53	-0.31	0.29	-0.13	-1.00	-1.00	-1.00	-0.65	0.63	0.54	0.09	-0.42	-1.00	0.17	-1.00	-0.16	-1.00	0.08	0.00	-0.16
23	Chemical engineering	0.10	-0.11	-0.19	0.02	0.04	-0.53	-0.19	0.10	-0.04	-0.06	0.04	0.06	0.03	-0.42	-0.27	0.05	0.05	-0.09	-0.04	0.03	0.18	-0.21	-0.29	-0.19	-0.07	-0.09	0.04
24	Environmental technology	0.00	-1.00	0.43	0.04	0.02	-1.00	-0.34	0.49	-0.04	0.00	-0.01	-0.12	-0.28	-1.00	0.06	0.26	0.12	-0.06	0.05	0.36	0.37	-0.18	-0.73	0.18	-0.13	-0.15	0.04
25	Handling	-0.15	0.16	-0.30	-0.01	0.00	-0.13	-0.34	0.06	0.17	-0.10	0.39	0.05	0.22	-0.43	-0.04	-0.16	0.41	-0.20	0.13	-0.08	0.03	0.39	-0.29	0.03	0.04	-0.17	-0.06
26	Machine tools	-0.39	-0.18	-0.04	-0.31	0.14	-0.70	-0.40	0.44	-0.13	-0.17	0.18	-0.40	-0.04	-0.36	-0.18	-0.40	0.41	-0.47	0.22	-0.28	0.01	-0.45	-0.40	-0.12	-0.25	0.11	-0.23
27	Engines, pumps, turbines	-0.33	0.15	-0.19	0.18	0.12	-0.25	-0.49	-0.09	-0.14	0.03	-0.04	-0.25	0.29	-0.12	0.35	-0.40	-0.60	-0.59	-0.12	-0.25	-0.35	0.23	-0.75	-0.21	-0.58	-0.19	0.06
28	Textile and paper machines	0.34	0.33	0.19	-0.27	0.10	-0.32	-0.56	-0.30	-0.21	-0.28	0.17	0.30	-0.39	-0.19	-0.31	-0.34	0.07	-0.15	0.06	-0.03	-0.25	-1.00	-0.46	-0.63	0.35	-0.33	-0.26
29	Other special machines	0.18	0.18	0.13	0.06	0.02	-0.28	-0.16	0.31	0.11	-0.07	0.22	-0.15	-0.35	0.05	0.29	0.21	-0.36	-0.06	0.11	-0.13	0.20	-0.54	-0.30	-0.01	-0.15	-0.15	-0.15
30	Thermal processes and apparatus	0.10	0.08	0.13	0.10	0.08	0.29	-0.14	-0.23	0.16	-0.13	0.24	-0.25	0.21	-1.00	0.26	-0.12	-0.50	-0.38	0.14	0.41	0.11	0.12	0.04	0.57	-0.14	-0.10	-0.22
31	Mechanical elements	-0.40	-0.23	-0.08	-0.16	0.17	-0.78	-0.50	-0.35	-0.14	-0.13	0.05	-0.39	-0.41	-1.00	-0.17	-0.12	-0.19	-0.57	0.04	-0.12	-0.19	-0.32	-0.48	0.03	-0.49	0.00	-0.11
32	Transport	-0.46	-0.35	0.31	-0.55	0.12	-0.28	-0.64	-0.39	0.10	0.11	0.07	-0.01	-1.00	-0.61	0.29	-0.43	-0.73	-0.57	-0.07	-0.14	-0.14	-0.43	-0.50	-0.01	-0.66	-0.02	-0.24
33	Furniture, games	0.03	0.14	-0.21	-0.06	-0.06	-0.59	0.05	-0.20	0.24	-0.14	0.36	0.27	-0.06	-1.00	-0.14	0.05	0.26	-0.11	0.35	0.09	0.36	0.37	0.26	-0.38	-0.42	-0.15	0.14
34	Other consumer goods	0.28	0.11	0.00	-0.27	-0.01	-0.61	-0.24	-0.18	0.08	0.01	0.31	0.45	-0.41	0.12	0.11	-0.15	0.65	-0.19	0.13	0.21	0.10	0.44	0.31	-0.22	-0.48	-0.30	0.00
35	Civil engineering	-0.04	-0.02	0.22	0.09	-0.03	0.02	-0.09	0.23	0.29	-0.07	0.25	0.03	-0.14	-0.13	0.14	0.01	0.63	-0.19	0.38	0.39	0.25	0.37	0.10	0.02	-0.21	-0.06	0.02

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Table 15: Rescaled RTAs for entrepreneurial ventures — per country & field

(Higher values indicate higher levels of specialisation)

FhG35	Field of technology	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	IT	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK
1	Electrical machinery, apparatus, energy	-0.38	-1.00	0.06	-0.28	0.05	-0.08	0.22	0.14	-0.25	-0.12	0.00	-1.00	-1.00	-1.00	-0.38	0.16	-0.20	0.07	0.00	-0.20	-0.30	-0.40	0.16	0.30	-0.20	-0.25	0.12
2	Audio-visual technology	-0.01	0.30	-0.37	0.08	-0.03	-0.21	-0.04	-0.56	-0.06	0.11	-0.29	-0.14	0.71	-1.00	0.43	0.14	-0.60	-0.04	-0.08	-0.37	0.46	-1.00	0.22	0.22	-0.09	0.11	0.10
3	Telecommunications	0.24	0.45	-0.06	-0.26	-0.23	0.11	0.36	-0.42	-0.13	-0.04	-0.36	0.43	-1.00	-1.00	-0.39	-0.60	-0.35	-0.20	-0.31	0.67	-0.36	-0.03	0.55	0.60	0.20	0.20	0.29
4	Digital communication	0.41	-1.00	-0.41	-0.47	-0.30	0.65	0.58	-0.59	0.06	0.15	-0.54	-1.00	-1.00	0.63	-0.23	-0.02	-0.37	-0.29	-0.38	-0.29	-1.00	0.56	0.48	0.48	0.34	0.00	0.31
5	Basic communication processes	0.66	-1.00	-1.00	-0.43	-0.19	-1.00	0.20	-1.00	-0.07	0.11	-0.62	-1.00	-1.00	-1.00	0.10	-1.00	-1.00	-0.42	-1.00	-1.00	-1.00	0.83	0.70	0.48	-0.15	0.41	0.24
6	Computer technology	-0.08	0.19	-0.47	-0.16	-0.19	-0.19	-0.04	-0.01	-0.42	0.23	-0.41	0.00	0.21	0.12	0.09	-0.09	-0.11	-0.16	-0.35	-0.16	0.02	0.01	0.05	0.43	0.31	0.29	0.24
7	IT methods for management	-0.11	0.34	0.21	-0.45	-0.22	0.62	0.24	-0.03	-0.13	-0.07	-0.30	0.25	-1.00	0.68	-0.14	0.27	-0.29	-0.29	-0.38	0.63	0.24	0.62	0.39	0.84	0.27	-0.09	0.32
8	Semiconductors	-0.39	-1.00	-0.23	-0.36	0.03	-1.00	-0.45	-0.26	-0.26	0.02	-0.40	0.21	0.33	-1.00	-0.53	-0.47	-1.00	-0.17	0.07	0.25	0.00	-1.00	-0.17	-1.00	0.24	-0.39	0.27
9	Optics	-0.44	0.05	-0.29	-0.45	-0.05	-0.12	-0.13	-0.08	-0.47	0.07	-0.42	0.57	0.28	0.87	0.82	0.12	-0.53	-0.05	0.01	-0.69	-0.26	-1.00	0.11	-1.00	0.18	0.25	0.11
10	Measurement	0.02	-1.00	-0.31	0.13	0.06	-0.54	-0.08	0.16	-0.31	-0.09	-0.14	-0.07	-1.00	0.63	-0.65	-0.24	-1.00	-0.17	-0.10	-0.01	-0.49	-0.38	0.25	-0.02	0.18	0.16	0.07
11	Analysis of biological materials	0.31	-1.00	-0.38	0.32	-0.09	0.58	-0.69	0.25	-0.22	-0.13	-0.43	0.05	-1.00	-1.00	-0.39	-0.13	-0.60	0.04	0.10	-0.38	-0.35	-1.00	-1.00	-0.32	0.26	0.44	0.07
12	Control	-0.18	-0.06	0.05	-0.18	-0.13	0.66	0.20	0.04	0.18	0.03	0.03	-0.47	0.16	-1.00	-0.65	0.39	-1.00	-0.08	-0.19	-0.27	0.08	-0.04	0.19	-1.00	0.03	-0.06	0.15
13	Medical technology	0.10	-0.63	-0.19	-0.21	0.05	-0.35	0.23	-0.01	-0.07	0.16	-0.06	0.05	-0.47	-0.25	0.35	0.10	-0.78	-0.12	-0.11	-0.14	-0.08	0.18	-1.00	-0.77	-0.28	-0.01	-0.03
14	Organic fine chemistry	-0.11	0.39	0.78	0.29	-0.06	0.53	-0.37	-0.17	0.12	0.05	-0.11	-0.03	0.69	-1.00	0.26	0.53	-0.03	-0.13	-0.35	-0.03	0.10	-1.00	0.32	-1.00	-0.23	0.02	0.02
15	Biotechnology	0.18	0.09	-0.54	0.38	-0.05	0.58	-1.00	-0.04	0.03	0.16	-0.48	0.42	0.31	-1.00	-1.00	-0.08	-1.00	0.19	0.04	-0.01	0.06	-1.00	0.01	-0.49	0.12	0.13	0.02
16	Pharmaceuticals	0.06	-0.07	0.11	0.50	-0.19	0.03	-0.02	-0.31	-0.17	0.03	-0.35	0.10	-1.00	-0.28	0.50	0.02	-0.79	-0.16	0.06	-0.34	0.09	-0.05	-0.25	-0.78	-0.19	0.27	0.13
17	Macromolecular chemistry, polymers	0.23	-1.00	0.52	-0.28	0.08	0.19	-0.65	0.59	-0.11	-0.34	-0.13	-1.00	-1.00	-1.00	-0.32	-0.25	-0.27	0.22	-0.12	-0.48	0.06	-1.00	-1.00	0.41	-0.40	-0.22	0.13
18	Food chemistry	0.35	0.29	-0.63	0.34	0.06	-0.23	-0.17	-0.41	0.13	0.10	-0.11	-1.00	0.17	0.65	-1.00	0.12	-1.00	0.49	-0.42	-0.15	0.09	-1.00	-1.00	-0.32	-0.38	-0.20	-0.29
19	Basic materials chemistry	0.11	0.56	0.21	0.04	0.04	0.42	-0.12	0.20	-0.08	-0.06	-0.19	-1.00	-1.00	-1.00	0.13	0.20	0.09	-0.01	-0.13	0.01	-0.36	-1.00	0.00	0.33	-0.20	-0.02	0.11
20	Materials, metallurgy	0.26	0.68	0.49	0.09	0.07	0.46	-0.77	-0.25	0.05	-0.01	-0.06	-1.00	0.61	-1.00	0.05	0.30	0.03	-0.52	0.00	0.44	0.02	0.46	-0.16	0.63	0.42	-0.14	-0.18
21	Surface technology, coating	-0.07	-1.00	-0.06	-0.28	0.02	-0.23	-0.84	0.16	-0.02	-0.05	-0.02	0.49	-1.00	-1.00	0.09	-0.60	0.73	-0.21	-0.26	0.37	-0.17	-0.03	0.20	-0.33	0.58	-0.09	-0.02
22	Micro-structure and nano-technology	-0.22	-1.00	-1.00	-0.45	-0.50	-1.00	0.18	0.63	-1.00	-0.39	-0.67	-1.00	-1.00	-1.00	-1.00	0.61	-1.00	0.22	-0.03	-1.00	-1.00	-1.00	-1.00	-1.00	0.57	0.70	0.25
23	Chemical engineering	0.03	-0.03	0.08	0.03	0.04	-0.21	-0.17	0.13	-0.14	-0.13	0.08	0.14	0.38	0.10	-0.37	-0.20	-0.22	-0.06	0.06	0.00	-0.07	-0.34	-0.31	-0.11	0.16	0.20	-0.13
24	Environmental technology	0.00	-1.00	0.40	0.06	0.03	-1.00	0.11	0.57	-0.05	-0.14	-0.01	-1.00	0.14	-1.00	-0.42	0.18	0.16	-0.05	0.25	0.25	-0.20	-0.06	-1.00	0.31	0.19	-0.03	-0.13
25	Handling	-0.08	0.01	-0.32	-0.07	0.00	-1.00	-0.14	-0.15	0.11	-0.05	0.23	-0.41	0.24	-0.06	-0.33	-0.19	0.20	0.09	-0.02	-0.16	0.21	0.37	0.27	-0.26	-0.28	-0.24	-0.14
26	Machine tools	-0.38	-1.00	-0.36	-0.26	0.17	-0.51	-0.41	-0.27	-0.18	-0.14	0.20	-0.27	0.38	-1.00	-0.55	-0.43	0.44	0.00	-0.01	-0.44	0.00	-1.00	-0.12	-1.00	0.13	0.05	-0.35
27	Engines, pumps, turbines	-0.07	0.56	0.03	-0.38	0.01	0.19	0.21	0.32	0.03	-0.15	0.03	-0.08	0.63	0.30	0.06	0.09	-0.68	0.07	0.02	-0.11	0.06	0.50	0.08	0.33	-0.34	-0.05	0.05
28	Textile and paper machines	-0.23	-1.00	-1.00	0.00	0.01	0.14	-0.45	-0.55	0.14	-0.36	0.33	0.62	0.20	-1.00	-0.49	-0.58	0.22	-0.12	0.03	-0.53	-0.60	-1.00	-1.00	-1.00	0.14	-0.15	-0.20
29	Other special machines	0.01	0.00	0.08	-0.09	0.03	-0.36	-0.11	0.20	0.07	0.04	0.12	-0.19	-1.00	-1.00	0.05	0.02	-0.57	0.23	-0.03	-0.34	0.23	-0.48	-0.03	0.06	-0.02	-0.28	-0.18
30	Thermal processes and apparatus	-0.36	0.15	0.30	-0.29	0.09	0.31	0.56	-0.16	0.06	-0.10	0.13	-0.57	0.36	-1.00	-0.15	-0.13	-0.69	-0.13	0.29	0.33	0.11	0.17	-0.45	0.31	-0.05	-0.23	-0.30
31	Mechanical elements	-0.30	0.17	-0.10	-0.10	0.10	-1.00	-0.45	-0.07	0.01	-0.18	0.15	-0.68	-1.00	-1.00	-0.89	-0.20	-0.05	-0.07	0.07	0.24	-0.23	0.00	0.14	0.30	-0.28	-0.19	-0.08
32	Transport	-0.18	0.19	-0.29	-0.42	0.01	-0.59	-0.33	-0.08	0.20	0.02	0.18	0.14	-1.00	-0.02	-0.43	-0.09	-0.67	0.01	0.18	0.16	-0.07	-0.13	-0.10	-0.09	-0.19	-0.28	-0.10
33	Furniture, games	-0.12	0.04	-0.23	-0.13	0.00	-0.60	0.05	-0.38	0.16	-0.07	0.11	-0.74	-0.27	-1.00	-0.92	0.06	-0.01	0.13	0.01	-0.21	0.18	0.07	0.24	-0.66	-0.34	-0.08	0.00
34	Other consumer goods	-0.12	-0.34	0.19	-0.45	-0.03	-0.49	0.04	-0.14	0.17	0.07	0.21	0.17	-0.13	0.12	-0.54	-0.09	0.61	0.02	-0.13	-0.43	0.13	0.22	0.16	-0.56	-0.46	-0.14	-0.11
35	Civil engineering	0.11	-0.07	-0.10	-0.28	0.04	-0.14	0.04	-0.01	0.20	0.00	0.11	0.14	-0.49	0.07	-0.58	0.03	0.50	0.02	0.17	0.13	-0.07	0.23	-0.24	-0.46	-0.06	-0.39	-0.16

Source: PATSTAT autumn 2011 edition, Amadeus 2012, internet searches based on applicant name.

Table 16: Rescaled RTAs for large companies — per country & field

(Higher values indicate higher levels of specialisation)

FhG35	Field of technology	BE	BG	CZ	DK	DE	EE	IE	EL	ES	FR	IT	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK
1	Electrical machinery, apparatus, energy	-0.30	-1.00	0.23	-0.20	0.06	-1.00	-0.86	-1.00	-0.11	0.03	-0.03	-1.00	-1.00	-1.00	-0.04	-0.57	-0.56	0.07	0.21	-0.07	-0.65	-1.00	-0.60	-1.00	-0.22	-0.50	-0.15
2	Audio-visual technology	-0.30	-1.00	-0.80	0.27	-0.26	-1.00	0.40	-1.00	-0.47	-0.03	-0.66	-0.04	-1.00	-1.00	-0.54	-1.00	-1.00	0.51	-0.14	-0.43	-0.41	0.22	-0.87	-1.00	0.05	-0.08	-0.05
3	Telecommunications	-0.78	-1.00	-0.79	-0.65	-0.38	-1.00	0.04	-1.00	0.07	0.18	-0.42	0.20	-1.00	-1.00	-0.74	-1.00	-1.00	0.02	-0.59	-0.40	-0.38	-1.00	-0.93	-1.00	0.62	0.55	0.01
4	Digital communication	-0.88	-1.00	-0.83	-0.77	-0.36	-1.00	0.20	-1.00	-0.36	0.28	-0.43	-1.00	-1.00	-1.00	-0.90	-1.00	-1.00	-0.02	-0.77	-0.79	-1.00	-1.00	-0.88	-1.00	0.61	0.50	-0.02
5	Basic communication processes	-0.61	-1.00	0.61	-0.39	-0.06	-1.00	-0.59	-1.00	-0.68	-0.07	-0.76	-1.00	-1.00	-1.00	-0.38	-1.00	-1.00	0.33	0.09	-0.50	-1.00	-1.00	-0.70	-1.00	0.24	0.14	-0.12
6	Computer technology	-0.38	-1.00	-0.51	-0.51	-0.11	-1.00	0.54	-1.00	-0.58	0.03	-0.48	0.55	-1.00	-1.00	-0.32	-1.00	-0.57	0.23	-0.37	-0.44	0.01	0.21	-0.93	-1.00	0.37	0.10	0.10
7	IT methods for management	-0.78	-1.00	-0.26	-0.47	-0.07	-1.00	0.86	-1.00	-0.07	0.07	-0.49	-1.00	-1.00	-1.00	-0.10	-1.00	-1.00	-0.15	0.05	0.18	0.64	-1.00	-1.00	-1.00	0.28	0.04	0.17
8	Semiconductors	-0.47	-1.00	-0.75	-0.71	0.11	-1.00	-0.67	-1.00	-0.68	-0.34	-0.62	-0.24	-1.00	-1.00	-0.02	-1.00	0.50	0.36	0.40	-0.70	-1.00	-1.00	-1.00	-1.00	-0.65	-0.73	-0.44
9	Optics	0.18	-1.00	-0.27	-0.50	-0.09	-1.00	-0.42	0.03	-0.60	-0.26	-0.46	-1.00	-1.00	0.91	-0.32	-1.00	-1.00	0.51	-0.30	-0.17	-1.00	-1.00	-0.36	-1.00	-0.34	-0.33	-0.10
10	Measurement	-0.35	-1.00	-0.68	-0.25	0.04	-1.00	-0.56	-1.00	-0.32	-0.09	-0.22	-0.29	-1.00	0.81	0.15	-1.00	0.71	0.14	-0.11	0.11	-0.34	0.76	-0.55	-1.00	-0.18	-0.15	0.11
11	Analysis of biological materials	0.39	-1.00	-1.00	0.35	-0.04	-1.00	0.19	-1.00	0.05	-0.07	-0.36	-1.00	-1.00	-1.00	-0.32	-1.00	-1.00	0.06	-0.20	-1.00	-1.00	0.86	-0.04	-1.00	-0.37	-0.10	0.23
12	Control	-0.52	-1.00	0.38	-0.42	0.08	-1.00	0.07	-1.00	0.03	0.05	-0.17	0.60	-1.00	-1.00	-0.21	-1.00	-0.10	-0.19	0.24	-0.27	0.20	0.39	-1.00	-1.00	-0.22	-0.06	-0.04
13	Medical technology	-0.49	0.81	0.02	0.50	-0.04	-1.00	0.55	-0.19	-0.28	-0.39	-0.03	0.33	-1.00	-1.00	-1.00	-0.73	-0.35	0.24	0.05	0.07	-1.00	-1.00	-0.11	-1.00	-0.57	0.17	0.08
14	Organic fine chemistry	0.21	-1.00	0.17	0.00	0.01	0.73	-0.24	0.15	0.33	0.21	0.00	0.09	0.90	-1.00	-1.00	0.80	-0.02	-0.19	-0.51	0.23	0.67	-1.00	0.65	0.63	-0.68	-0.59	0.08
15	Biotechnology	0.51	-1.00	0.25	0.75	-0.05	0.95	-0.02	-1.00	-0.09	-0.22	-0.43	-1.00	-1.00	-1.00	-0.88	0.56	-1.00	0.00	-0.14	-1.00	-0.05	-1.00	0.39	0.63	-0.63	-0.28	0.02
16	Pharmaceuticals	0.45	0.87	-0.26	0.45	-0.16	0.66	0.49	0.75	0.66	-0.07	0.18	-0.23	0.78	-1.00	-0.85	0.85	0.64	-0.27	0.04	0.36	0.66	0.02	0.85	0.47	-0.83	0.23	0.24
17	Macromolecular chemistry, polymers	0.48	-1.00	-0.61	-0.49	0.06	-1.00	-0.05	-1.00	-0.59	-0.11	0.24	-1.00	-1.00	-1.00	-0.93	-1.00	-1.00	0.04	-0.29	-1.00	-0.62	-1.00	-0.85	0.07	0.09	-0.78	-0.32
18	Food chemistry	0.41	-1.00	-1.00	0.69	-0.25	-1.00	-0.20	0.63	-0.05	-0.24	-0.08	-1.00	-1.00	-1.00	-0.84	0.01	-1.00	0.44	-0.56	-0.46	-1.00	-1.00	-1.00	-1.00	-0.56	-0.87	0.17
19	Basic materials chemistry	0.25	-1.00	-0.66	0.08	0.08	-1.00	-0.30	-1.00	-0.29	-0.21	-0.07	-1.00	-1.00	-1.00	-0.52	-0.18	-1.00	0.06	-0.27	-0.45	-0.67	-1.00	-0.87	0.59	-0.48	-0.71	0.15
20	Materials, metallurgy	0.31	-1.00	-0.27	-0.16	0.03	-1.00	-0.53	-1.00	-0.21	0.01	-0.10	0.24	-1.00	-1.00	0.69	-1.00	-1.00	-0.30	0.45	0.30	-0.16	-1.00	-1.00	0.76	-0.19	-0.08	0.00
21	Surface technology, coating	0.27	-1.00	-0.75	-0.32	0.06	-1.00	-0.37	0.03	-0.32	-0.03	-0.01	-1.00	-1.00	-1.00	-0.06	-1.00	-1.00	-0.15	0.10	-0.17	-0.56	0.73	-0.36	-1.00	-0.11	-0.08	0.00
22	Micro-structure and nano-technology	-0.48	-1.00	-1.00	-0.80	0.07	-1.00	-0.42	-1.00	-1.00	-0.09	-0.34	-1.00	-1.00	-1.00	-0.09	-1.00	-1.00	0.25	-0.38	-1.00	-1.00	-1.00	0.33	-1.00	0.06	-0.68	-0.11
23	Chemical engineering	0.18	-1.00	-0.37	0.03	0.06	-1.00	-0.56	-0.41	-0.27	-0.13	0.01	-1.00	-1.00	-1.00	-0.36	-0.68	0.56	-0.11	-0.23	0.05	0.13	0.24	-0.64	-1.00	-0.10	-0.12	0.13
24	Environmental technology	0.07	-1.00	0.26	-0.07	0.04	-1.00	-0.88	-1.00	-0.22	0.00	-0.10	-1.00	-1.00	-1.00	0.32	-1.00	-1.00	-0.04	-0.21	-1.00	-0.37	-1.00	-1.00	-1.00	-0.18	-0.19	0.13
25	Handling	-0.16	-1.00	-0.28	-0.04	0.02	-1.00	-0.51	-1.00	0.16	-0.04	0.44	-0.35	-1.00	-1.00	-0.19	-1.00	0.64	-0.29	0.17	0.10	-0.40	-1.00	-0.74	-1.00	0.10	-0.19	-0.01
26	Machine tools	-0.51	-1.00	0.05	-0.43	0.14	-1.00	-0.43	0.89	-0.22	-0.19	0.09	0.02	-1.00	-1.00	-0.08	-1.00	-0.23	-0.54	0.34	-0.15	0.08	0.28	-0.43	0.07	-0.32	0.15	-0.13
27	Engines, pumps, turbines	-0.34	-1.00	-0.10	0.20	0.12	-1.00	-1.00	-1.00	-0.35	0.08	0.00	-1.00	-1.00	-1.00	0.55	-1.00	-1.00	-0.71	-0.15	-0.67	-0.73	-1.00	-0.90	-1.00	-0.66	-0.30	0.13
28	Textile and paper machines	0.47	-1.00	0.50	-0.38	0.05	-1.00	-0.47	-1.00	-0.42	-0.26	0.24	-1.00	-1.00	-1.00	-0.61	-1.00	-0.05	-0.14	0.15	0.04	-0.19	-1.00	-0.49	0.25	0.38	-0.33	-0.26
29	Other special machines	0.29	-1.00	0.23	-0.02	0.04	-1.00	-0.37	0.07	0.05	-0.05	0.27	-1.00	-1.00	-1.00	0.58	-0.70	-0.30	-0.16	0.03	-0.45	-0.25	-1.00	-0.76	-1.00	-0.25	-0.10	-0.18
30	Thermal processes and apparatus	0.26	-1.00	-0.22	0.17	0.10	-1.00	-0.75	-1.00	0.08	-0.10	0.30	-1.00	-1.00	-1.00	0.55	-1.00	-1.00	-0.51	-0.04	0.61	0.34	-1.00	0.05	-1.00	-0.17	-0.18	-0.31
31	Mechanical elements	-0.38	-1.00	-0.04	-0.16	0.16	-1.00	-0.81	-1.00	-0.17	-0.07	0.06	-0.21	-1.00	-1.00	-0.12	-0.77	-0.44	-0.67	0.00	-0.57	0.14	-1.00	-0.74	-0.16	-0.56	-0.06	-0.04
32	Transport	-0.52	-1.00	0.56	-0.69	0.12	-1.00	-0.96	-1.00	0.30	0.20	0.08	0.38	-1.00	-1.00	0.28	-0.84	-0.57	-0.75	-0.22	-0.52	-0.43	-1.00	-0.76	0.55	-0.77	-0.03	-0.25
33	Furniture, games	0.11	-1.00	-0.07	-0.30	-0.02	-1.00	0.19	-1.00	0.07	-0.02	0.45	-1.00	-1.00	-1.00	-0.47	-1.00	-1.00	-0.12	0.49	0.36	-0.09	0.52	0.12	-1.00	-0.49	-0.33	0.11
34	Other consumer goods	0.39	-1.00	-0.20	-0.31	0.03	-1.00	-0.63	-1.00	-0.07	0.09	0.36	0.77	-1.00	-1.00	0.31	-1.00	-1.00	-0.27	0.22	0.68	-1.00	-1.00	0.37	-1.00	-0.61	-0.52	-0.07
35	Civil engineering	-0.12	-1.00	0.39	0.24	-0.02	-1.00	-0.07	-1.00	0.26	0.01	0.25	-0.30	-1.00	-1.00	0.23	-0.65	-1.00	-0.24	0.41	0.71	0.61	-1.00	0.03	0.41	-0.22	0.06	0.01

Source: PATSTAT autumn 2011 edition, Amadeus 2012, Internet searches based on applicant name

Larger innovators such as Germany, the UK and France show less explicit technological specialisation than the smaller economies in the sample. Given their size, these Member States are simply more likely to allocate resources to a broader spectrum of technologies.

Next, full NIS RTAs for each field of technology *t* were regressed (OLS) on the RTAs resulting from incumbent and entrepreneurial innovative activity:

$$RTA_{NIS,t} = \alpha + \beta_1 RTA_{Large,t} + \beta_2 RTA_{Entrepreneurial,t} + \epsilon$$

In order to avoid distortion caused by RTAs stemming from NISs with few patents, countries filing fewer than 200 patent applications (all fields) were excluded from the analysis ([21]). Robust standard errors were estimated. **Table 17** provides an overview of the technologies where the RTA of entrepreneurial ventures and/or incumbent firms coincides with overall country-level RTAs.

Table 17: Technological specialisation per firm size and specialisation per country (Member States with at least 200 patent applications)

Significant correlation with large-firm specialisation	Significant correlation with entrepreneurial specialisation: no	Significant correlation with entrepreneurial specialisation: yes
no		Analysis of biological materials [11] (-, *) Micro-structure and nano-technology [22] (-, **) Environmental technology [24] (-, ***)
yes	Semiconductors [8] (***, -) Measurement [10] (**, -] Biotechnology [15] (*, -) Macromolecular chemistry, polymers [17] (***, -) Food chemistry [18] (***, -) Materials, metallurgy [20] (***, -) Surface technology, coating [21] (***, -) Chemical engineering [23] (***, -) Engines, pumps, turbines [27] (***, -) Other special machines [29] (***, -) Mechanical elements [31] (***, -) Transport [32] (***, -) Furniture, games [33] (***, -) Civil engineering [35] (***, -)	Electrical machinery, apparatus, energy [1] (***, ***) Audio-visual technology [2] (***, ***) Telecommunications [3] (***, **) Digital communication [4] (***, **) Basic communication processes [5] (***, *) Computer technology [6] (***, **) IT methods for management [7] (***, ***) Optics [9] (***, **) Control [12] (***, ***) Medical technology [13] (***, ***) Organic fine chemistry [14] (***, ***) Pharmaceuticals [16] (***, **) Basic materials chemistry [19] (***, *) Handling [25] (***, **) Machine tools [26] (***, ***) Textile and paper machines [28] (***, ***) Thermal processes and apparatus [30] (***, ***) Other consumer goods [34] (***, ***)

— p>0.1; * p<0.1; ** p<0.05; *** p<0.01.

Source: PATSTAT autumn 2011 edition, Amadeus 2012, Internet searches based on applicant name.

The results in **Table 17** suggest that nationwide specialisation levels are driven by large companies' specialisation patterns. Three technologies appear to be exclusively determined by entrepreneurial specialisation: Analysis of Biological Materials [11], Micro-structure and Nano-technology [22], and

([21]) In ascending order: Lithuania, Romania, Cyprus, Bulgaria, Latvia, Hungary, Slovakia, Greece, Malta, Estonia and Portugal.

Environmental Technology [24]. For a selection of technologies, entrepreneurial ventures and large firms complement each other in shaping the nationwide technological specialisation profile. In the area of Electrical Engineering, this is the case for all fields but one (Semiconductors [8]). In other areas, more heterogeneous patterns are observed: for Instruments, small-firm specialisation in Optics [9], Control [12] and Medical Technology [13] correlates significantly with overall specialisation. Among the chemistry-related technologies, incumbent and entrepreneurial specialisation in Organic Fine Chemistry [14], Pharmaceuticals [16] and Basic Materials Chemistry [19] closely reflect overall specialisation. In the area of Mechanical Engineering, a similar pattern holds for Handling [25], Machine Tools [26], Textile and Paper Machines [28], and Thermal Processes and Apparatus [30].

As data are standardised at country level (so that technologies can be compared across NISs), estimators should be interpreted accordingly. In other words, we report 'average' patterns, in which each country is included as one observation, without adjusting for the size of the countries in question.

In interpreting the above results, we should avoid statements about causality. Due to the use of cross-sectional data, estimators can at best point to the existence of a relationship between components of the NIS and its performance as a whole. To test whether, and to what extent, entrepreneurial and incumbent firms are responsible for RTAs would require longitudinal data and the use of panel data estimation techniques. In addition, a more extensive dataset would enable the inclusion in the model equation of additional RTAs measuring specialisation in other patenting sectors (e.g. universities).

4. Conclusions and suggestions for future research

The current study considers how SMEs' involvement in inventive activity in the EU can be measured on the basis of patents. The growing interest in innovation has produced a broad range of patent indicators, but indicators of SMEs' contribution have thus far been lacking. This comes as no surprise, given a number of challenges in terms of data treatment and assembly. The main part of this paper (Section 2) presents a methodology for estimating the proportion of all patents filed by EU companies accounted for by SMEs. This methodology allows for the large-scale identification of SMEs among patent applicants. In an additional analytical part (Section 3), the classified patent portfolios are used to shed light on the SME footprint in different areas of technology.

A first step to obtaining reliable indicators of SMEs' share of EU corporate patenting activity consists of matching firms' patent data to financial data (*matching*). Next, multiple matches are disambiguated (*disambiguation*) and — using available financial and ownership information — firm size is determined for the resulting unique matches (*classification*). For a considerable proportion of corporate applicants, no match can be found in financial directories or, if a match is found, information on entity size and, above all, dependency status is lacking. Thus, automated matching and disambiguation procedures need to be complemented with additional efforts in order to obtain an accurate estimate of SMEs' share of corporate patent activity (*extrapolation*).

Combining automated matching to financial directories with additional searches gives us precise estimators of SMEs' share of patent activity. For the EU as a whole, we find that 79 % of all patent technology can be attributed to large firms and 17 % to SMEs. For 4 %, the size of the corporate applicant remains unclear. At the same time, SMEs' contribution varies considerably across Member States.

Additional analysis (Section 3) focusing on SMEs' contribution in different areas of technology identifies the comparative advantages for the 16 technologically most active Member States. The results signal a distinctive contribution from SMEs in a considerable number of technological fields. Using multiple regression analysis per field of technology, SME specialisation patterns (RTAs) were related to overall national specialisation patterns. For 21 of the 35 fields of technology, there is a significant correlation between SME specialisation and national specialisation. While this SME contribution to national specialisation patterns is (in a majority of fields) complemented by large firms' contribution as well, specialisation seems to be spearheaded by SMEs in a number of emerging fields, including Environmental Technology, Analysis of Biological Materials, and Micro-structure and Nano-technology. These findings underline the intertwining of SME and large-firm technological development in the EU's industrial landscape.

In terms of methodology, the contribution of our findings is to demonstrate the feasibility of producing SME patent indicators. At the same time, there is clear potential for future improvements. The coverage of financial directories is crucial in this respect: to the extent that these directories can incorporate more information on size and, especially, on the dependency status of firms, refining and maintaining efforts to create this type of indicator will become less cumbersome and time consuming.

5. References

Publications

Audretsch, D. B. (1995). *Innovation and industry evolution*, Cambridge, Massachusetts: MIT Press.

Baumol, W. J. (2002). *The free-market innovation machine: analysing the growth miracle of capitalism*. Princeton, New Jersey: Princeton University Press.

Baumol, W. J. (2004). *Entrepreneurial enterprises, large established firms and other components of the free-market growth machine*, Small Business Economics 23: pp. 9-21.

Belenzon, S. and Berkovitz, T. (2007). *Innovation in business groups*, CEP Discussion Paper no. 833.

Bessen, J. (2009). *Matching patent data to Compustat firms*, NBER PDP Project user documentation. http://www.nber.org/~jbessen/matchdoc.pdf

BvDEP Ownership Database (2008).

CHI Research (2003). *Small serial innovators: the small firm contribution to technical change*, for SBA Office of Advocacy (contract no. SBAHQ-01-C-0149).

Cochran, W. G. (1977). *Sampling techniques*, third edition. Wiley.

Cohen, W. H. 'Fifty years of empirical studies of innovative activity and performance'. In: Hall. B. H. and Rosenberg, N. (ed.) (2010). *Handbook of the economics of innovation, volume 1*: Elsevier.

EC (2005), *The new SME definition: user guide and model declaration* (2005). Enterprise and industry publications, European Commission.

Eurostat (2011), *Patent Statistics at Eurostat: Methods for Regionalisation, Sector Allocation and Name Harmonisation*, 2011 edition.

Eurostat (2011). *Key figures on European business* (with a special feature on SMEs), 2011 edition. Eurostat Pocketbooks.

Frietsch, R., Neuhaeusler, P., and Rothengatter, O. (?). *SME Patenting — An empirical analysis in nine countries*. Fraunhofer Institute for systems and innovation research ISI.

Iversen, E.J., Mäkinen, I., Lööf, H., Oh, D.-H., Jespersen, S.T., Junge, M. and Bech, J. (2009). *Small Nordic Enterprises, Developing IPR in Global Competition*, Norden: Nordic Innovation Centre study.

Hall, B.H., Jaffe, A. and Trajtenberg, M. (2001). *The NBER patent citations data file: lessons, insights and methodological tools*, NBER working paper series 8498.

Hall, B.H., Jaffe, A. and Trajtenberg, M. (2005). *Market value and patent citations*. The RAND Journal of Economics 36 (1): pp. 16-38.

Helmers, C. and Rogers, M. (2009). *Patents, entrepreneurship and performance*. Global COE Hi-Stat Discussion Paper Series 095.

Jensen, P.H. and Webster, E. (2006). *Firm size and the use of intellectual property rights*, The Economic Records, 82 (256): pp. 44-55.

Keupp, M. M., Lhuillery, S., Garcia-Torres, M.A., and Raffo, J. (2009). SME-IP — 2nd report. *Economic Focus Study on SMEs and intellectual property in Switzerland*. Swiss Federal Institute of Intellectual Property. Publication No 5 (06.09).

Macartney, G.J., under supervision of: Griffith, R. (2007). *Market institutions and firm behaviour: employment and innovation in the face of reform*, PhD Economics University College London.

OECD (2000). *Background report for 'Enhancing the competitiveness of SMEs in the global economy: strategies and policies. Workshop 1: enhancing the competitiveness of SMEs through innovation.'* Conference for ministers responsible for SMEs and industry ministers, Bologna, Italy, 14-15 June 2000.

Perrin, H. and Speck, K. (2004). *SMEs as patent applicants, The dossiers from the Observatory for Intellectual Property (INPI)*, December 2004.

Rothwell, R. (1989). *Small firms, innovation and industrial change*. Small Business Economics 1(1): pp. 51-64.

Scherer, F. M. (1991). 'Changing Perspectives on the Firm Size Problem', in Z. J. Acs and D. B. Audretsch (eds.), *Innovation and Technological Change. An International Comparison*, Ann Arbor: University of Michigan Press.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Cambridge, Massachusetts: Harvard University Press.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*, 2nd ed., New York, New York: Harper and Brothers.

Squicciarini, M., and Dernis, H. (2012). *A cross-country characterisation of the patenting behavior of firms based on matched firm and patent data*. Working paper: OECD working Party on Industry Analysis (DSTI/EAS/IND/WPIA(2012)5.

Thoma, G., Torrisi, S., Gambardella, A., Guellec, D., Hall, B.H. and Harhoff, D. (2010). *Harmonising and combining large datasets: an application to firm-level patent and accounting data*, OECD Directorate for Science, Technology and Industry working paper 2010/99.

Websites

Company Registration Office (Ireland) website:
http://www.cro.ie/search/CompanySearch.aspx

Kruispuntbank der Ondernemingen (Belgium) website:
http://kbopub.economie.fgov.be/kbopub/zoeknaamfonetischform.html

Annex 1: sampling methodology

The sample size is calculated by means of stratified random sampling. The sample size n_D in the full target population D is calculated as:

$$n_D = \frac{(\sum_{h=1}^{H} W_h . S_h)^2}{\mathrm{V}(\hat{\theta}_D) + \frac{1}{N_D}\sum_{h=1}^{H} W_h . S_h^2}$$

where $\mathrm{V}(\hat{\theta}_D)$ is the variance of the estimated overall share of SMEs, H is the number of strata in the target population D, $W_h = N_h / N_D$, where N_h is the number of enterprises in stratum h; N_D is the total number of enterprises in target population D; and S_h^2 is the stratum variance for the SME dummy variable y_a.

The numerator reflects the variance within each stratum multiplied by the size of the strata. In other words, for strata with more firms, more n will be included. A similar observation can be made with respect to the variance (S_h); this will be highest when the proportion of large and small firms is equal (50 x 50 = 2 500 whereas 90 x 10 = 900).

The stratum variance, S_h^2 can be expressed as follows:

$$S_h^2 = \frac{1}{N_h - 1}\sum_{k \epsilon a_h}\left(y_a(k) - \frac{1}{N_h}\sum_{k \epsilon a_h} y_{a(k)}\right)^2$$

In practice, the stratum variance S_h is not known. The variance per country per quantile for the matched applicants with sufficient information to determine company size is used as a proxy. To calculate the stratum variance for the SME dummy variable for strata reporting fewer than 10 % SMEs among the matched applicants with sufficient information, the SME percentage was set to 10 %, to ensure that at least some firms were sampled.([22])

The confidence interval for the estimated overall proportion of SMEs, with approximate confidence level of 95 %, is given by:

$$\hat{\theta}_D \pm 1{,}96.\sqrt{V(\hat{\theta}_D)}$$

The precision, α_D (set at 0.025 for a two-sided alternative) in terms of the length of the confidence interval:

$$\propto_D = 1{,}96.\sqrt{V(\hat{\theta}_D)}$$

From which it can be deduced that the variance $V(\hat{\theta}_D)$ can be expressed as:

$$V(\hat{\theta}_D) = \left(\frac{\propto_D}{1{,}96}\right)^2$$

([22]) Also, an additional sub-classification of SMEs into micro, small and medium-sized enterprises would result in higher variance and therefore require a bigger sample to obtain reliable estimates.

n_D Aiming for greater precision here will result in higher values for since the variance parameter, which is affected by the level of precision, is squared in the denominator (with a value < 1).

It is assumed that all strata are equally important and hence the *Neymann allocation* (Cochran 1977) can be used. The total sample size in the target population is distributed among strata, so the sample size in stratum h, n_h is given by:

$$n_h = n_D . \frac{N_h . S_h}{\sum_{h=1}^{H} N_h . S_h}$$

Decimals resulting from strata sample size computation are rounded up to the next integer. In addition, due to the skewed patent volume distribution — a minority of companies tend to account for more than half of the patent volume in most countries — the minimum sample size for the, on average smaller, top quantiles with populations of 200 applicants or fewer is set at 5 ([23]). The resulting sample sizes per stratum, and the population values on which their computation is based, are reported in **Table 8**. Strata with 200 applicants or fewer account for 2 952 of the total population of applicants. The calculated sample sizes for strata containing more than 200 applicants represent 72 804 applicants or the rest of the population. In total, 1 849 applicants have to be verified: 433 applicants represent strata containing no more than 200 applicants, 1 416 applicants account for the remaining strata with more than 200 applicants.

To illustrate the sampling methodology, the computation of the third stratum for non-matched Belgian corporate applicants containing 669 patentees is explained. Sequentially, the computation of the parts constituting the formula for the stratum sample size n_h is illustrated.

The proportion of the stratum population in the full target population is calculated as follows:

$$W_{str.3\backslash NM\backslash BE} = \frac{N_{str.3\backslash NM\backslash BE}}{N_D} = \frac{669}{75\,567} = 0.0089$$

As a proxy for S_h^2 — the stratum variance for the SME dummy variable y_a — the variance per country per quantile for the matched applicants with sufficient information to determine company size is used. In the case of the third stratum for Belgium (669 corporate applicants), matching Amadeus with PATSTAT resulted in the identification of 213 SMEs and 359 large companies.

$$S^2_{str.3\backslash NM\backslash BE} = \frac{1}{(213+359)-1} * \begin{bmatrix} 210 * \left(1 - \frac{1}{(213+359)}(213*1+359*0)\right)^2 + \\ 359 * \left(0 - \frac{1}{(213+359)}(213*1+359*0)\right)^2 \end{bmatrix} = 0.2341$$

Departing from a required 5% significance level for the proportion of SMEs, the α_D is set at 0.025 against a two-sided alternative:

$$V(\hat{\theta}_D) = \left(\frac{0.025}{1.96}\right)^2 = 0.000163$$

([23]) With a 200-observation population threshold per stratum, a sample size was calculated for the following strata only, using the full sample size calculation methodology specified in Cochran (1977):

- for the population of matched applicants with insufficient financial data, the third stratum for Austria, Belgium, Germany, Denmark, Spain, Finland, France, the UK, Ireland, Italy, the Netherlands and Sweden, and the second stratum for Spain, the UK and Italy;
- for the non-matched applicants, the third stratum for Austria, Belgium, Germany, Spain, Finland, France, the UK, Hungary, Italy, Luxemburg, the Netherlands and Sweden, and the second stratum for Spain and Italy.

For the remaining strata, a sample of five was taken where the population of the stratum was greater than five.

The full sample size n_D for all strata with populations of 200 or more (see above) is then computed as ([24]):

$$n_D = \frac{\left[\begin{matrix}(W_{str.i\backslash MBU\backslash AT} * S_{str.i\backslash MBU\backslash AT}) + \cdots + (W_{str.i\backslash MBU\backslash SE} * S_{str.i\backslash MBU\backslash SE}) + \\ (W_{str.i\backslash NM\backslash AT} * S_{str.i\backslash NM\backslash AT}) + \cdots + (W_{str.i\backslash NM\backslash SE} * S_{str.i\backslash NM\backslash SE})\end{matrix}\right]^2}{0.000163 + \frac{1}{75\,567} * \left[\begin{matrix}(W_{str.i\backslash MBU\backslash AT} * {S_{str.i\backslash MBU\backslash AT}}^2) + \cdots + (W_{str.i\backslash MBU\backslash SK} * {S_{str.i\backslash MBU\backslash SE}}^2) + \\ (W_{str.i\backslash NM\backslash AT} * {S_{str.i\backslash NM\backslash AT}}^2) + \cdots + (W_{str.i\backslash MBU\backslash SK} * {S_{str.i\backslash NM\backslash SE}}^2)\end{matrix}\right]} = 1{,}399$$

with *i* representing the quantile number of the strata with populations of more than 200 applicants.

Finally, to compute the sample size that is representative for the third stratum of unmatched Belgian corporate applicants, the following formula is solved:

$$n_h = n_D . \frac{\frac{669}{75\,567} * 0.2341}{\left[\begin{matrix}(W_{str.i\backslash MBU\backslash AT} * S_{str.i\backslash MBU\backslash AT}) + \cdots + (W_{str.i\backslash MBU\backslash SK} * S_{str.3\backslash MBU\backslash SE}) + \\ (W_{str.i\backslash NM\backslash AT} * S_{str.i\backslash NM\backslash AT}) + \cdots + (W_{str.i\backslash NM\backslash SK} * S_{str.3\backslash NM\backslash SE})\end{matrix}\right]} = 13$$

([24]) Due to rounding, the sum of the sample sizes for strata with more than 200 applicants in Tables 9 and 10 is higher (1 416).

Annex 2

Table 18: Results of the first round of classification of EU-27 corporate applicants based on financial firm-size indicators only

Country	Large				Small				No financial size indicators		Not matched		Total
	Certainly		Most likely		Certainly		Most likely						
	#	%	#	%	#	%	#	%	#	%	#	%	#
EU-27	8 850	8.5	455	0.4	28 342	27.2	26 320	25.3	529	0.5	39 670	38.1	104 166
BE	238	10.7	12	0.5	905	40.8	386	17.4	1	0.0	676	30.5	2 218
BG	20	18.7		0.0	16	15.0	9	8.4		0.0	62	57.9	107
CZ	68	13.6	3	0.6	214	42.8	51	10.2		0.0	164	32.8	500
DK	197	5.5	6	0.2	929	25.9	963	26.8	6	0.2	1 492	41.5	3 593
DE	3 184	10.6	175	0.6	4 421	14.7	8 524	28.3	16	0.1	13 810	45.8	30 130
EE	2	1.8	2	1.8	41	36.6	20	17.9		0.0	47	42.0	112
IE	54	4.4	9	0.7	181	14.7	639	51.7	29	2.3	323	26.2	1 235
EL	8	3.8		0.0	34	16.3	14	6.7	3	1.4	150	71.8	209
ES	350	8.3	3	0.1	1 733	40.9	309	7.3		0.0	1 839	43.4	4 234
FR	999	9.3	35	0.3	3 782	35.1	758	7.0	13	0.1	5 176	48.1	10 763
IT	1 195	9.1	24	0.2	5 505	42.0	2 247	17.1	3	0.0	4 130	31.5	13 104
CY	1	0.4	1	0.4	12	4.9	16	6.5	32	13.1	183	74.7	245
LV	3	4.1		0.0	9	12.2	6	8.1		0.0	56	75.7	74
LT		0.0		0.0	4	25.0	4	25.0		0.0	8	50.0	16
LU	29	4.5	17	2.6	73	11.2	140	21.6		0.0	390	60.1	649
HU	15	2.9		0.0	107	20.9	57	11.1	2	0.4	332	64.7	513
MT	4	4.9		0.0	7	8.5	42	51.2		0.0	29	35.4	82
NL	388	5.6	24	0.3	3 143	45.6	1 110	16.1	55	0.8	2 171	31.5	6 891
AT	288	9.5	6	0.2	293	9.6	1 044	34.3	1	0.0	1 410	46.4	3 042
PL	56	14.0	7	1.7	92	22.9	83	20.7		0.0	163	40.6	401
PT	34	8.9	3	0.8	129	33.8	26	6.8		0.0	190	49.7	382
RO	4	7.0		0.0	13	22.8		0.0		0.0	40	70.2	57
SI	15	5.7		0.0	14	5.3	46	17.4	60	22.6	130	49.1	265
SK	13	10.5		0.0	49	39.5	14	11.3		0.0	48	38.7	124
FI	224	8.3	9	0.3	751	28.0	737	27.5	3	0.1	959	35.7	2 683
SE	320	5.1	9	0.1	2 822	45.3	296	4.8	5	0.1	2 774	44.6	6 226
UK	1 141	7.0	110	0.7	3 063	18.8	8 779	53.8	300	1.8	2 918	17.9	16 311

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Table 19: Overall matching and size classification results for corporate applicants (%)

Country	Large	Small — large group	Small — maj. owned by institutional investor	Small — maj. owned by public body	Small — small group	Small & independent	Small — maj. owned by individual	Small with insufficient ownership info	Matched but insuff. fin. and ownership info	Not matched	Total
EU-27	**8.9**	**5.4**	**0.5**	**0.2**	**1.9**	**11.1**	**9.7**	**23.8**	**0.5**	**38.1**	**100.0**
BE	11.3	6.7	0.3	0.0	0.9	8.7	0.9	40.6	0.0	30.5	100.0
BG	18.7	3.7	0.0	0.0	0.0	3.7	5.6	10.3	0.0	57.9	100.0
CZ	14.2	2.6	0.0	0.0	0.2	12.2	18.4	19.6	0.0	32.8	100.0
DK	5.6	4.3	1.4	0.1	3.3	13.8	2.6	27.2	0.2	41.5	100.0
DE	11.1	5.9	0.2	0.1	0.4	9.6	12.7	14.1	0.1	45.8	100.0
EE	3.6	1.8	1.8	0.0	0.0	24.1	8.9	17.9	0.0	42.0	100.0
IE	5.1	4.0	0.3	0.0	1.7	3.5	1.1	55.8	2.3	26.2	100.0
EL	3.8	0.5	0.0	0.0	0.0	6.7	8.6	7.2	1.4	71.8	100.0
ES	8.3	3.8	0.6	0.0	0.7	6.6	9.1	27.3	0.0	43.4	100.0
FR	9.6	5.7	0.9	0.0	0.7	6.3	4.0	24.5	0.1	48.1	100.0
IT	9.3	4.1	0.3	0.0	1.0	21.5	16.5	15.7	0.0	31.5	100.0
CY	0.8	1.2	0.0	0.0	0.4	2.0	1.6	6.5	12.7	74.7	100.0
LV	4.1	2.7	0.0	0.0	0.0	2.7	1.4	13.5	0.0	75.7	100.0
LT	0.0	6.3	0.0	0.0	0.0	12.5	0.0	31.3	0.0	50.0	100.0
LU	7.1	2.8	0.0	0.0	0.2	6.0	6.2	17.7	0.0	60.1	100.0
HU	2.9	0.2	0.0	0.0	0.0	1.8	0.6	29.4	0.4	64.7	100.0
MT	4.9	1.2	1.2	0.0	1.2	6.1	1.2	48.8	0.0	35.4	100.0
NL	6.0	5.6	1.0	1.7	4.1	2.9	0.5	46.0	0.7	31.5	100.0
AT	9.7	4.7	0.3	1.5	0.1	9.4	10.6	17.5	0.0	46.4	100.0
PL	15.7	2.5	0.2	0.0	0.2	12.7	12.5	15.5	0.0	40.6	100.0
PT	9.7	3.4	0.0	0.0	0.8	13.6	6.8	16.0	0.0	49.7	100.0
RO	7.0	1.8	0.0	0.0	0.0	1.8	12.3	7.0	0.0	70.2	100.0
SI	5.7	1.1	0.4	0.0	0.0	4.5	7.2	9.8	22.3	49.1	100.0
SK	10.5	2.4	0.0	0.0	0.0	8.1	4.8	35.5	0.0	38.7	100.0
FI	8.7	4.2	0.2	0.1	4.1	7.2	4.0	35.6	0.1	35.7	100.0
SE	5.3	4.9	0.3	0.0	5.5	2.5	1.1	35.7	0.1	44.6	100.0
UK	7.7	6.9	0.8	0.2	4.1	18.6	14.3	27.9	1.7	17.9	100.0

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Table 20: Overall matching and 'large' vs. SME classification results for corporate applicants

(%)

Country	'Large' company	SME	Matched but unknown	Not matched	Total
EU-27	**15.0**	**13.0**	**33.9**	**38.1**	**100.0**
BE	18.3	9.6	41.6	30.5	100.0
BG	22.4	3.7	15.9	57.9	100.0
CZ	16.8	12.4	38.0	32.8	100.0
DK	11.4	17.1	30.0	41.5	100.0
DE	17.3	10.0	26.8	45.8	100.0
EE	7.1	24.1	26.8	42.0	100.0
IE	9.5	5.2	59.2	26.2	100.0
EL	4.3	6.7	17.2	71.8	100.0
ES	12.9	7.3	36.4	43.4	100.0
FR	16.3	7.1	28.6	48.1	100.0
IT	13.7	22.5	32.3	31.5	100.0
CY	2.0	2.4	20.8	74.7	100.0
LV	6.8	2.7	14.9	75.7	100.0
LT	6.3	12.5	31.3	50.0	100.0
LU	9.9	6.2	23.9	60.1	100.0
HU	3.1	1.8	30.4	64.7	100.0
MT	7.3	7.3	50.0	35.4	100.0
NL	14.3	7.0	47.2	31.5	100.0
AT	16.1	9.5	28.0	46.4	100.0
PL	18.5	13.0	27.9	40.6	100.0
PT	13.1	14.4	22.8	49.7	100.0
RO	8.8	1.8	19.3	70.2	100.0
SI	7.2	4.5	39.2	49.1	100.0
SK	12.9	8.1	40.3	38.7	100.0
FI	13.2	11.3	39.8	35.7	100.0
SE	10.6	8.0	36.9	44.6	100.0
UK	15.5	22.7	43.8	17.9	100.0

Source: PATSTAT autumn 2011 edition, Amadeus 2012.

Table 21: Extrapolation to population: corporate applicants (%)

Country	'Large'	Unknown — non-matched	Unknown — matched	SME
EU-27	**33.8**	**3.2**	**8.9**	**54.0**
BE	48.5	9.3	0.0	42.2
BG	38.6	0.0	14.9	46.5
CZ	41.4	0.0	0.2	58.3
DK	25.3	4.6	13.6	56.4
DE	40.8	6.3	13.0	39.9
EE	20.9	1.8	0.0	77.3
IE	19.1	0.8	8.0	72.1
EL	31.5	0.0	35.8	32.7
ES	35.4	0.0	6.1	58.6
FR	41.6	0.5	9.6	48.3
IT	24.1	0.4	6.6	68.9
CY	8.2	0.0	22.0	69.8
LV	10.3	0.0	16.2	73.4
LT	42.9	0.0	0.0	57.1
LU	27.1	0.0	35.1	37.9
HU	33.7	0.0	12.5	53.9
MT	22.7	0.0	6.6	70.7
NL	28.2	3.0	5.2	63.6
AT	45.7	0.0	6.9	47.4
PL	56.3	7.9	0.0	35.8
PT	27.9	3.0	11.5	57.6
RO	50.5	1.8	0.0	47.7
SI	20.5	0.0	9.1	70.4
SK	37.0	0.0	8.1	54.9
FI	32.8	3.6	10.7	52.9
SE	31.2	0.8	9.3	58.6
UK	26.2	3.6	4.0	66.2

Source: PATSTAT autumn 2011 edition, Amadeus 2012, Internet searches based on applicant name.

Glossary

BvD	Bureau Van Dijk
EC	European Commission
EPO	European Patent Office
EU	European Union
EUR	euro
Eurostat	Statistical Office of the European Union
FhG35	Fraunhofer 35 technology classification
Fin.	financial
FTE	full-time equivalent (staff headcount)
ICT	information and communications technology
IT	information technology
Maj.	majority
NIS	national innovation system
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Leased Squares
PATSTAT	EPO Worldwide Patent Statistical Database
PCT	Patent Cooperation Treaty
R&D	research & development
RTA	relative technological advantage
SME	micro, small or medium-sized enterprise
USPTO	United States Patent and Trademark Office
WIPO	World Intellectual Property Organisation

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions:

- via one of the sales agents of the Publications Office of the European Union (http://publications.europa.eu/others/agents/index_en.htm).

KS-GQ-14-009-EN-N

Publications Office

doi:10.2785/55102