ANNEX A

Methodological Issues in the Development of Indicators of Innovation and Transfer in Environmental Technologies

by
Nick Johnstone, Ivan Haščič and Fleur Watson (OECD Environment Directorate)

Introduction[1]

This book presents analyses of the determinants of environment-related innovation – assessing the relative importance of different factors (policy measures, market conditions, scientific capacity, etc.) on the rate and pattern of innovation and the diffusion of these inventions in the world economy. Both of these strands of work depend on the availability of appropriate indicators. This annex reviews the methodological aspects of development of such indicators. It is intended to serve as a reference document for papers arising out of this work.

The indicators of innovation in environment-related technologies (ENV-tech) thus complement those indicators which have earlier been developed by the OECD Science, Technology and Industry Directorate in the areas of Information and Communications Technologies, Biotechnology, or Nanotechnology (see OECD *Patent Statistics Manual*, 2009).[2]

The annex is organised as follows. The next section reviews the different indicators commonly used to measure innovation. This is followed by a review of alternative indicators of technology transfer. In both cases it is argued that in the context of environment-related technologies, patent data offer a good alternative to the existing measures. The final section then presents the methodology of the development of indicators of innovation and transfer based on patent data.

Indicators of innovation in environmental technologies

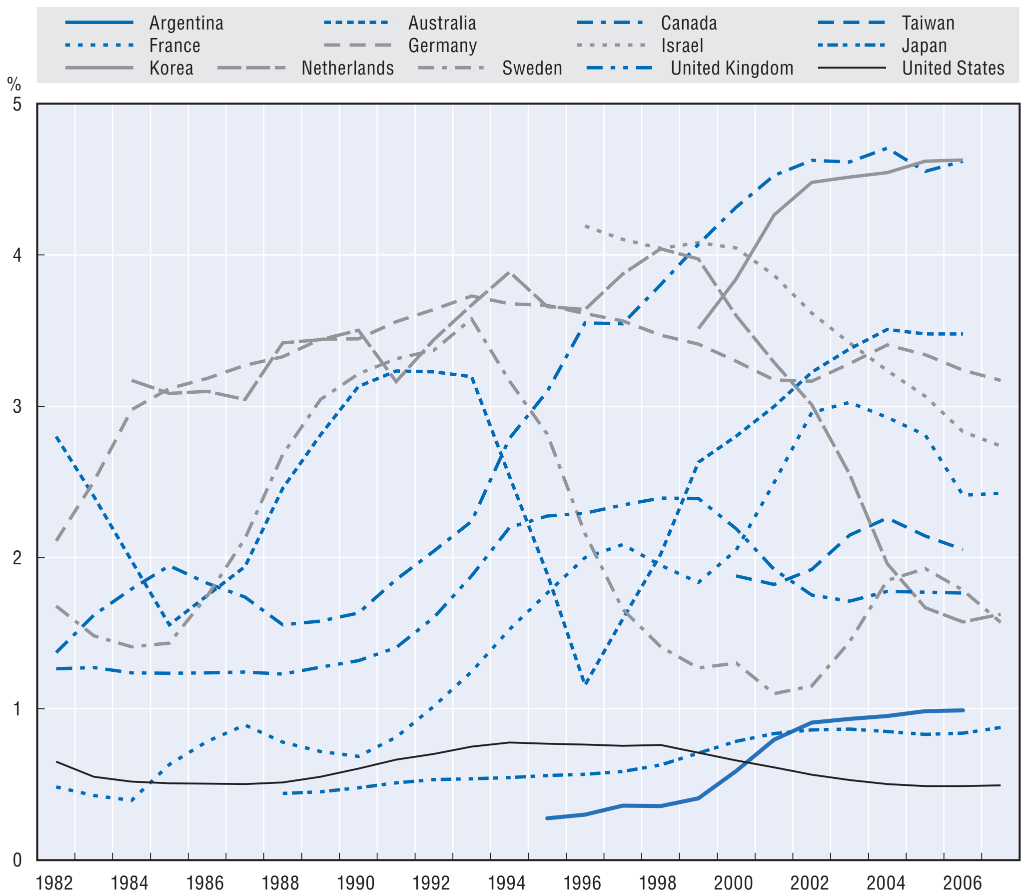
There are a number of candidates for the measurement of innovation (see OECD *Main Science and Technology Indicators*, 2008a). Most commonly, ***R&D expenditures or the number of scientific personnel*** in different sectors are frequently used as indicators. However, a sub-set of OECD and non-OECD countries have also begun to collect data on *government budget appropriations and outlays for R&D* (GBAORD) by socio-economic objective, including "control and care for the environment".[3] Figure A.1 provides some evidence for countries for which this data is relatively complete in recent years. For most countries, the data indicate that between 0.5% and 4% of GBAORD is specifically targeted at environmental objectives. While in large economies such as Germany, Japan and the US this share has remained relatively stable, there seems to have been a large degree of variation across countries and over time.

However, there has been much greater uniformity among countries in GBAORD expenditures directed at "rational utilisation of energy". For most countries, the data indicate that up to 5% of GBAORD is specifically targeted at energy objectives, although this share was higher in the early 1980s (Figure A.2).

Finally, data are also available on government R&D spending in the energy sector (see *OECD/IEA Energy Technology R&D Statistics*). The data include expenditures directed at energy generation from fossil fuels, nuclear energy, and renewable energy, hydrogen and fuel cells, energy storage, as well as measures directed at improving energy efficiency in industry, residential and commercial uses, and transportation. Figure A.3 gives the proportion of total energy technology R&D directed at renewable energy and energy

Figure A.1. **Percentage of GBAORD expenditures directed at "control and care for the environment"**

3-year moving average



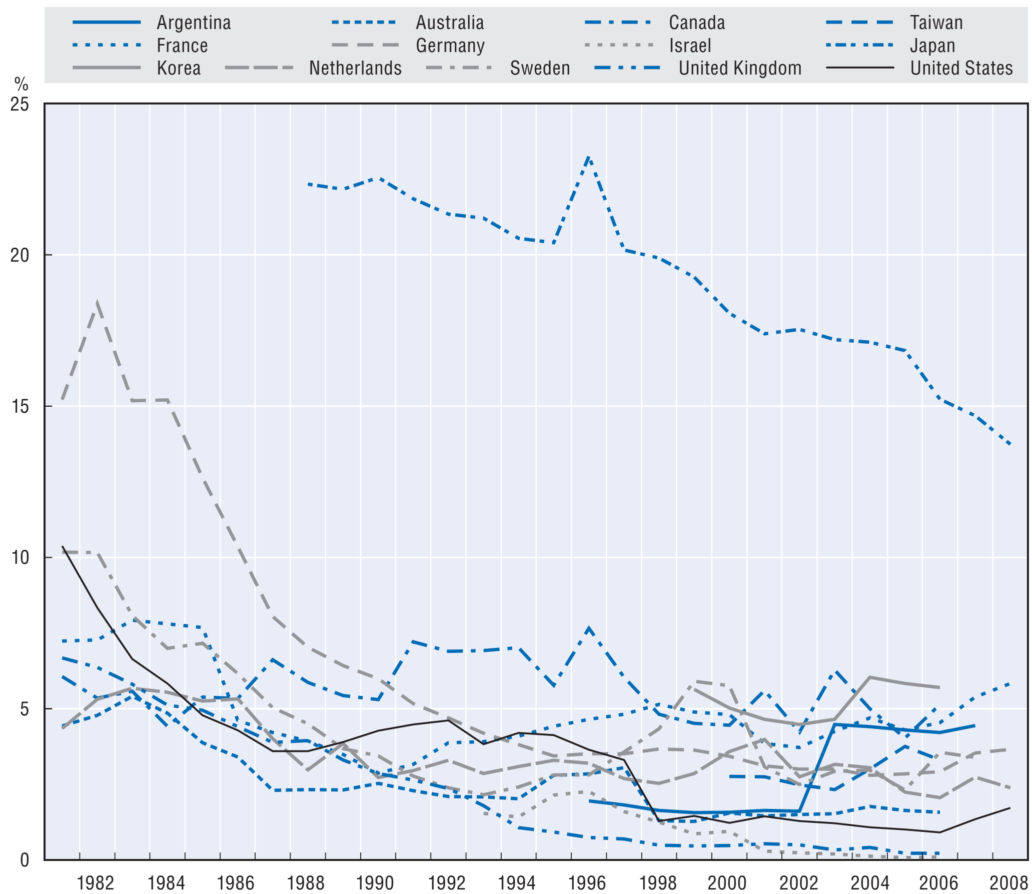
Source: OECD.Stat (*www.oecd.org/statistics*), *Research and Development Statistics.*

efficiency measures. On the one hand, countries such as Finland, Denmark, Sweden, Austria, Ireland and Hungary, devote a relatively high share of their budgets to such aims. On the other hand, this share is much lower for countries such as France, Japan, Germany, Australia, Korea, Canada and the US, often as a consequence of higher levels of spending on nuclear energy sources.

Although such indicators do reflect an important element of the overall innovation system, there are a number of disadvantages associated with their use as indicators of innovation. For example, with respect to private R&D expenditures, the data are incomplete. Further, the data are only available at an aggregate level and (with the exception of the energy sector) they cannot be broken down by technology group. Moreover, there is no source of data for private R&D expenditures by socio-economic objective that is comparable to that used for GBAORD. Perhaps most significantly, R&D expenditures are measures of inputs to the innovation process, whereas an "output" measure of innovation would be broadly preferable.

Given the general lack of data in this area, several ***micro-level data*** collection efforts have, therefore, been undertaken which have sought to measure innovation outputs. For instance, in the European Union, a small number of "environment-related" questions have

Figure A.2. **Percentage of GBAORD expenditures directed at "rational utilisation of energy"**



Source: OECD.Stat (*www.oecd.org/statistics*), *Research and Development Statistics*.

been applied as part of the *Community Innovation Survey*. Figure A.4 gives the respondents' perceptions from that *Survey* of the importance of the effects of their innovation efforts. It is noticeable that "environmental" factors rank at the bottom of this list.

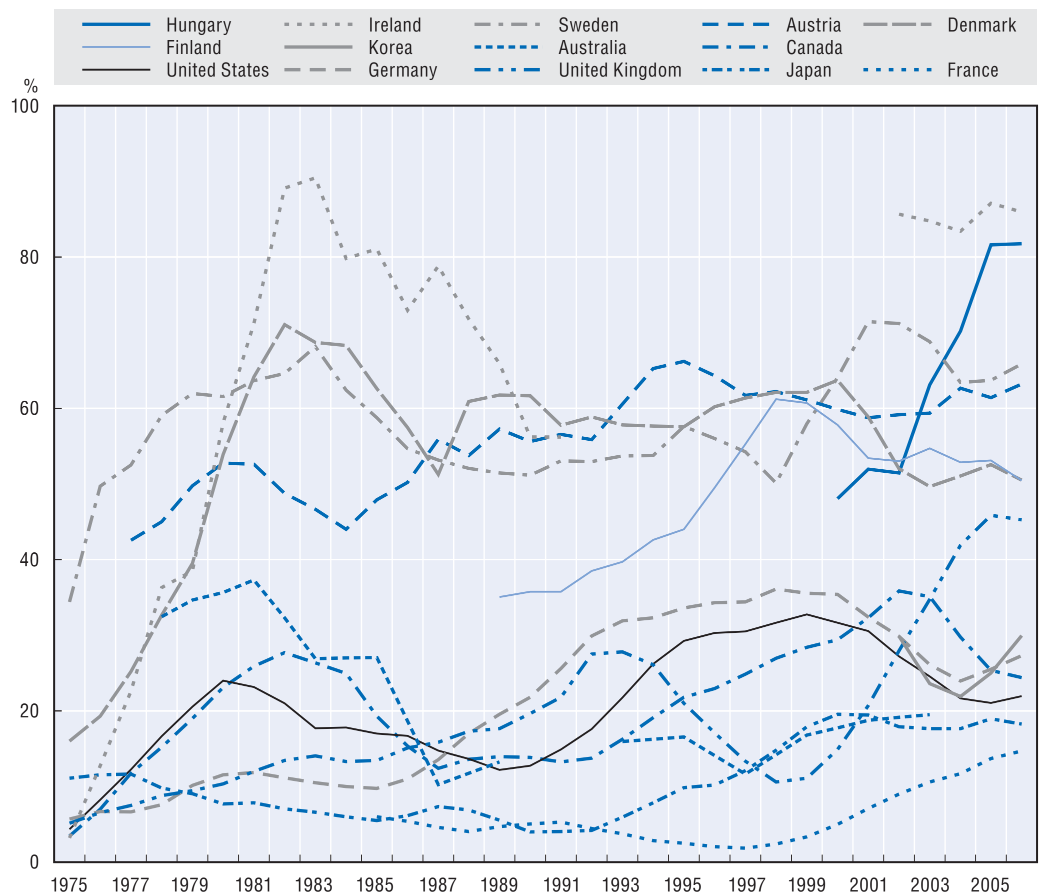
While "environmental" concerns were only addressed tangentially in previous rounds of the CIS, considerable effort has gone into the design of the most recent CIS survey questionnaire, in order to ensure that environmental concerns are addressed in a much more systematic manner in the future (see Box A.1).

Several of the researchers involved in the design of the environmental components of the CIS questionnaire were also involved in the (2006) OECD project on "Environmental Policy and Firm-Level Management" (Johnstone, 2007). In this latter project, data was collected on *input* measures of "environmental innovation", such as expenditures on environment-related R&D, as well as on *output* measures such as "clean production" and "product design". For illustration, Figure A.5 provides data on the percentage of firms in that project (by industrial sector) which reported having taken environmental factors into account in *product design*.

The main shortcoming with such exercises is their cost. A dedicated industrial survey which addresses environmental concerns on a periodic basis would be prohibitively

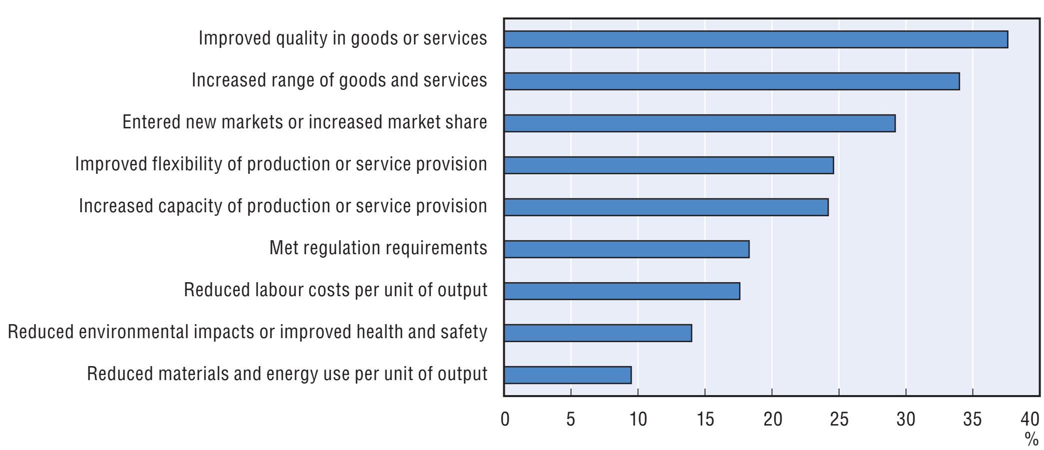
Figure A.3. **Percentage of energy technology R&D expenditures directed towards "renewable energy" and "energy efficiency" measures**

3-year moving average



Source: OECD.Stat (*www.oecd.org/statistics*), *Energy Technology R&D Statistics*.

Figure A.4. **Motivations identified as highly important for innovation activities (CIS 4-EU27)**



Source: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-07-113/EN/KS-SF-07-113-EN.PDF.

Box A.1. "Environmental" section of CIS 5

Innovations with environmental benefits

An environmental innovation is a new or significantly improved product (good or service), process, organisational method or marketing method that creates environmental benefits compared to alternatives.

- The environmental benefits can be the primary objective of the innovation or the result of other innovation objectives.
- The environmental benefits of an innovation can occur during the production of a good or service, or during the after sales use of a good or service by the customer.

1. During the three years 2006 to 2008, did your enterprise introduce a product (good or service), process, organisational or marketing innovation with any of the following environmental benefits?

Environmental benefits from the production of goods or services within your enterprise:

Reduced material use per unit output	❒ Yes ❒ No
Reduced energy use per unit output	❒ Yes ❒ No
Reduced CO_2 footprint for your enterprise	❒ Yes ❒ No
Replaced materials with less polluting or hazardous substitutes	❒ Yes ❒ No
Reduced soil, water, or air pollution	❒ Yes ❒ No
Recycled waste, water, or materials	❒ Yes ❒ No

Environmental benefits from the after sales use of a good or service by the customer:

Reduced energy use	❒ Yes ❒ No
Reduced air, water, soil or noise pollution	❒ Yes ❒ No
Improved recycling of product after use	❒ Yes ❒ No

2. During 2006-08, did your enterprise introduce an environmental innovation in response to:

Need to comply with existing environmental regulations	❒ Yes ❒ No
Environmental regulations that you expected to be introduced in the future	❒ Yes ❒ No
Availability of government grants, subsidies or other financial incentives for environmental innovation	❒ Yes ❒ No
Market demand from your customers for environmental innovations	❒ Yes ❒ No
Voluntary codes for environmental good practice within your sector	❒ Yes ❒ No

3. Does your enterprise have procedures in place to regularly identify and reduce your enterprise's environmental impacts? (For example preparing environmental audits, setting environmental performance goals, ISO 14001 certification, etc.)

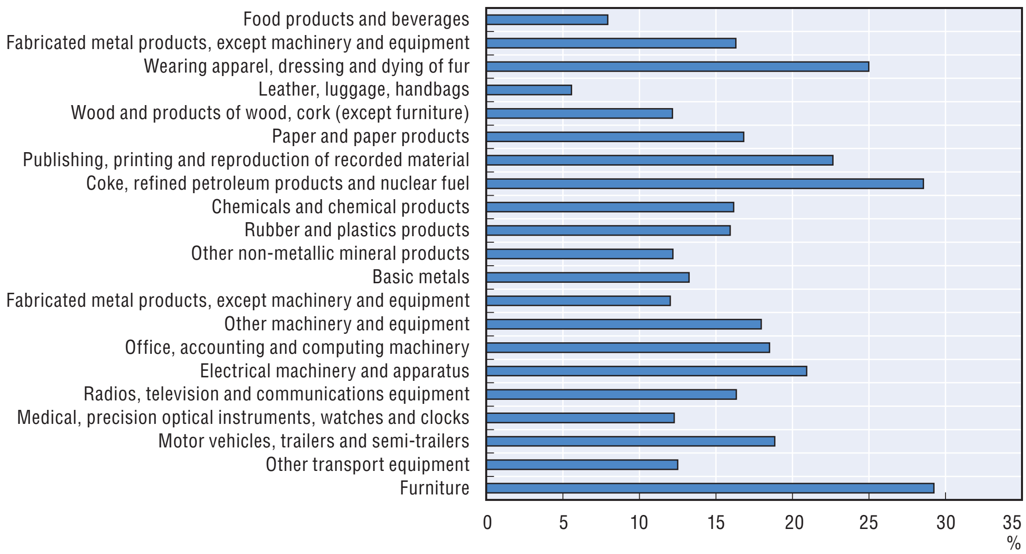
❒ Yes: implemented before January 2006.

❒ Yes: Implemented or significantly improved after January 2006.

❒ No.

expensive. While some countries do have "environmental" components in their standard industrial censuses or innovation surveys (*for example*, Canada, Norway, Japan), these data are not comparable across countries, and therefore cannot be used to develop indicators across countries.

Figure A.5. **Percentage of firms which report having taken environmental factors into account in product design**



Source: Johnstone (2007).

It is therefore necessary to look elsewhere for sources of information/data which can be used to derive indicators. One possibility would, of course, be the development of indicators based upon existing ***sectoral and commodity classifications*** – which have been developed to measure the output of goods and services. To the extent that new technologies are contained in direct (embodied) form in goods and services that are produced, such forms of innovation would be reflected in the base data. However, this would first require identification of sector or commodity classifications which represent "environmental" technologies.

The OECD/Eurostat Informal Working Group on the Environment Industry has developed a Manual (OECD, 1999) which provides a framework for the definition and classification of "environmental industry activities" (see Sinclair-Desgagne, 2008, for a recent discussion). This *Manual* identifies three broad "environmental segments", each of which includes a large range of business activities:

- Pollution management, including goods that help control air pollution, manage wastewater and solid waste, clean up soil, surface water and groundwater, reduce noise and vibrations, and facilitate environmental monitoring, analysis and assessment.
- Cleaner technologies and products, including goods that are intrinsically cleaner or more resource-efficient than available alternatives. For example, a solar photovoltaic power plant is cleaner than a coal-fired one.
- Resource management, including goods used to control indoor pollution, supply water, or help to manage farms, forests or fisheries sustainably. Also included are goods used to conserve energy and goods that help prevent or reduce environmental impacts of natural disasters, such as fire-fighting equipment.

However, and as pointed out in the Manual itself, standard sectoral codes (for example, ISIC, NACE, NAICS) do ***not*** lend themselves to such a breakdown, except in very specific areas such as water supply, wastewater treatment, and solid waste treatment and disposal.

Moreover, such categories relate primarily to "end-of-pipe" solutions to environmental concerns, areas where innovation is likely to be increasingly less beneficial overall.

On the basis of commodity classifications (the Harmonised Commodity Description and Coding System), the OECD has developed an illustrative list of "environmental goods" (see OECD, 2001) – broken down into the following broad headings:

A) **Pollution management:**

- Air pollution control.
- Wastewater management.
- Solid waste management.
- Remediation and clean-up of soil and water.
- Noise and vibration abatement.
- Environmental monitoring analysis and assessment.

B) **Cleaner technologies and products:**

- Cleaner/resource-efficient technologies and processes.
- Cleaner/resource-efficient products.

C) **Resource management group:**

- Indoor air pollution control.
- Water supply.
- Recycled materials.
- Renewable energy plant.
- Heat/energy saving and management.
- Sustainable agriculture and fisheries.
- Sustainable forestry.
- Natural risk management.
- Eco-tourism.

This list has since informed discussions about tariff arrangements related to "environmental goods and services" at the World Trade Organisation (WTO), in the context of the Doha Round of multilateral trade negotiations – which calls *inter alia* for the liberalisation of trade in "environmental goods" (and services).

However, it is important to note that these headings do not feature in the Harmonised System. The commodity codes themselves refer to generic commodity classifications. Indeed, many of the codes included in the list encompass goods and services which have a range of uses besides environmental protection (see OECD 2007 for a discussion of this issue). For instance, the list includes "air compressors mounted on a wheeled chassis for towing" (8414.40) or "articles of cast iron" (7325.10). More significantly, "environmental" goods are often designated as such in relation to a conventional alternative, which may well be included in the very same commodity classification – *i.e.*"parts for spark-ignition internal combustion piston engines" (8409.91).

And finally, classification of a good as being "environmental" does not provide any particular indication of the amount of "innovation" it represents – although production of goods and services is an important determinant and consequence of innovation, clearly only a small percentage of production can be considered to constitute "technological innovation".

In sum, commodity classifications cannot be used to develop indicators of "environmental innovation", for two key reasons:

- The commodity classifications do not lend themselves to the identification of goods and services with beneficial environmental consequences. In most cases, the classes used are much broader than the intended "target", including goods which have no specific environmental implications. Worse, the classifications are sufficiently broad that they include goods which may well be the "dirty" substitutes for "environmental innovations".
- The commodity classifications do not allow for the distinction between standardised goods and services which have been on the market for some time, and those goods and services which represent real technological innovations.

Fortunately, there are other possible "output" indicators which address both of these concerns: bibliometric data (scientific publications) and technometric data (patent publications). The use of ***bibliometric data*** as a measure of innovation has been given renewed impetus with the growth of the Internet, combined with increasingly efficient search engines. Using keywords and indexing codes, searches of relevant databases (for example, the Science Citation Expanded Index) are typically undertaken here. Data on author, affiliation, date of publication, etc. can be extracted, and counts can be developed to assess the relative innovative activity (see Meyer, 2002).

This kind of indicator is particularly useful for analyzing the diffusion of knowledge among inventors (and between countries), based on co-publications and citations. However, there are also some shortcomings associated with the use of bibliometric data. In particular, while such data is indeed an "output" indicator of innovation, it is only an indirect indicator of a market output. Publication in a peer-reviewed journal reflects a scientific advance, but not necessarily one which has commercial applications. It is therefore difficult to use citations even as an index of quality, let alone of actual economic importance.

As an alternative, ***patent data*** have often been used as a measure of technological innovation because they focus on outputs of the inventive process (Griliches, 1990 and OECD, 2009). Patent data provide a wealth of information on the nature of the invention and the applicant, the data is readily available (if not always in a convenient format), discrete (and thus easily subject to the development of indicators). Significantly, there are very few examples of economically significant inventions which have not been patented (Dernis *et al.*, 2001).

Most importantly, the application-based nature of the patent classification systems allows for a richer characterisation of relevant technologies. Since the International Patent Classification (IPC) system includes over 70 000 separate classification codes, it is possible to identify very precise technological fields. In the area of "environment-related innovation", some examples of relevant codes include:

- *B60L 7/10: Electrodynamic brake systems for vehicles – dynamic electric regenerative braking for vehicle.*
- *C02F 3/28: Biological treatment of water, waste water or sewage using anaerobic digestion processes.*

- *F03D 3/02: Wind motors with rotation axis substantially at right angle to wind direction – having a plurality of rotors.*

Consequently, patent data ***can be disaggregated to specific technological areas***, as was done in previous OECD work in this area, which examined the cases of renewable energy, wastewater effluent and motor vehicle emissions (OECD 2008b). However, recent developments in the *EPO PATSTAT Database* have enabled the implementation of search strategies which provide broader coverage of the data than those previously available (OECD triadic patent family database and commercial providers). Most significantly, this data allow for the possibility to undertake much more refined and accurate searches of innovations in different spheres.

In the environmental sphere, this allows for the identification of more "integrated" technological innovations (which are virtually impossible to identify using the other data sources discussed earlier). However, it is important to recognise that patents cannot be used to develop a measure of all innovations. First, they are designed only to protect technological inventions. Other IPR regimes exist to protect innovations in other fields – for example, copyrights for literature, trademarks for words or graphic devices which distinguish a product and the registration of designs, and where protection is focussed on the appearance of a produc. In addition, less formal ways (than intellectual property rights) to protect technological inventions also exist – notably industrial secrecy, or purposefully complex technical specifications. Surveys of inventors indicate that the rate at which new innovations are patented varies across industries. A further critique of patent data relates to the fact that not everything that is patented is eventually commercialised and adopted. There are, however, significant fees attached to the examination of a patent application (and to renewal fees, once the patent has been granted). So it is safe to assume that, at least in the expectations of the applicant or patent holder, the prospects for commercialisation and adoption are good. Nonetheless, the economic value of patents varies. For meaningful empirical analyses, it is therefore important to control statistically for differences in the propensity to patent, the scope of the claims, the value of the patent and other factors which vary across countries, time and technology fields.

Indicators of international transfer of environmental technologies

Technology transfer can be either "embodied" or "disembodied", and take place either through market or non-market means. A possible taxonomy might take the following form (see Maskus, 2004):

- Market-mediated transfer:
  - Trade in goods and services.
  - Foreign direct investment.
  - Licensing.
  - Joint ventures.
  - Cross-border movement of personnel.
- Non-market transfer:
  - Imitation and reverse engineering.
  - Employee turnover.
  - Published information (journals, test data, patent applications).

Available empirical evidence strongly supports the finding that the bulk of technology transfer takes place via: **i) *trade;* ii) *foreign direct investment (FDI); and* iii) *licensing*** (Maskus, 2004). Precisely which channel is most important depends in part upon the characteristics of the "recipient country" (*i.e.* domestic research capacity, strength of intellectual property rights regimes, etc.) and the nature of the technology being transferred (*i.e.* the potential for imitation and reverse engineering).

When seeking to assess the spatial patterns and rates of international technology transfer, it is therefore important to focus on measures which reflect potential transfer through these primary channels. Since technologies may be transferred in direct (embodied) form through trade in goods and services, such forms of transfer would be reflected in ***trade*** data. However, this would necessitate the identification of relevant sector or commodity classifications which represent "environmental" technologies.

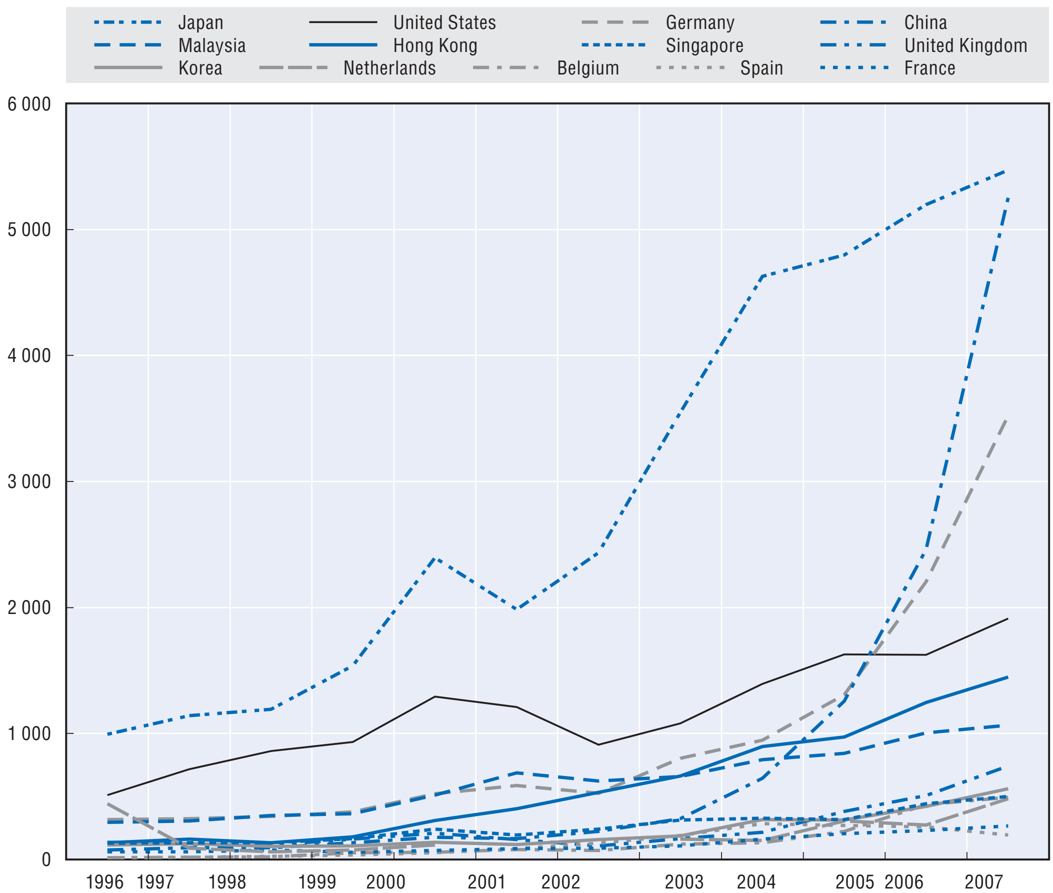
As noted above the OECD has developed an illustrative list of "environmental goods" (see OECD 2001). This list has since informed discussions about environmental goods and services at the World Trade Organisation (WTO), in the context of the Doha Round of multilateral trade negotiations – which calls for the liberalisation of "environmental goods" (and services). In principle, based upon this list, it is possible to examine recent trends in the export of "environmental goods and services".

For instance, according to figures published in OECD *Indicators of Globalisation* in 2006, exports of environmental goods in the OECD area reached USD 370 billion (1% of its GDP and nearly 6% of its merchandise exports). In the same year, BRICS countries (Brazil, the Russian Federation, India, China and South Africa) exported USD 43 billion, which accounted for almost 1% of their GDP and 2.7% of their total merchandise exports. Over the last four years, trade in "environmental goods" also pursued a dynamic pattern of growth, increasing faster than total merchandise trade particularly in the BRICS (where exports have been growing at an annual average rate of 35%). More than 25% of exports of "environmental goods" are for wastewater treatment equipment, which is also the fastest growing segment of the market. This is followed by air pollution control, waste management and environmental monitoring equipment.

However, and as noted above, the list of commodity classes used to extract the data used in these figures – while undoubtedly valuable for negotiating purposes – cannot be used credibly for statistical purposes. This is because a large number of the classes involved are only peripherally related to environmental concerns, and in some cases may even relate primarily to the "brown" substitutes for "green" alternatives (for example, "parts for spark-ignition internal combustion piston engines" – HS 8409.91).

The implications that this has for the assessment of the development of indicators of international technology transfer can be seen with reference to renewable power. HS 8541.40 is proposed as a measure of solar power technologies and Figure A.6 gives the trend in exports for the G7 countries, as well as for China and Spain. The remarkable level and growth rate of exports from China is certainly attributable to the breadth of the definition applied, which includes not only photovoltaic devices but also light-emitting diodes and semiconductor devices. Indeed, Hong Kong, China is the world's fifth largest exporter of this commodity class. Presumably, this is due to a high proportion of re-exports from other countries through Hong Kong, China. However, due to missing data on re-exports for most top exporting countries it is not possible to calculate *net* exports.

Figure A.6. **"Solar power" technology exports (based on COMTRADE data) (million USD)**



And finally, and of even more relevance to *this* paper, trade in an "environmental" good need not actually constitute "technology transfer" – although trade is an important channel of international technology transfer, not all trade can be considered to be technology transfer. In particular, trade in standardised goods and services can hardly be considered technology transfer. For several reasons, therefore, trade data is *not* an appropriate means by which to examine the transfer of environmental technologies.

Technology can also be transferred through ***foreign direct investment***. If a subsidiary of a multinational corporation is established, the parent company may transfer advanced technologies directly to the subsidiary. This may diffuse more widely in the economy by different channels – *for example*, local employees of the subsidiary taking up employment in domestic firms, and carrying knowledge about the technology with them. However, it is even more difficult in this case (than it was in the trade case examined above) to identify potential transfers which are directly relevant to environmental concerns. FDI data is not available at a level of disaggregation that would allow for an assessment of "environment-related" trends.

And finally, technologies can be transferred through ***explicit licensing of specific technologies***. However, data on licensing is very sparse, and to the authors' knowledge, no effort has yet been made to assess licensing in any sector (or with respect to any particular good or service) which might be considered to be "environmental".

The idea of using ***patent data*** to measure international technology transfers arises from the fact that there will be a partial "trace" of all three of the above-noted channels of transfer in patent applications. If there is any potential for reverse engineering, exporters, investors and licensors will each have an incentive to protect their intellectual property when it goes overseas. Although it cannot capture the full extent of the transfers which eventually take place, patent data *can* provide robust indicators of trends in both the direction and the extent of international transfer. Patent data has already been used extensively for this purpose, although not in the environmental sphere (see Eaton and Kortum, 1996 for the seminal study).

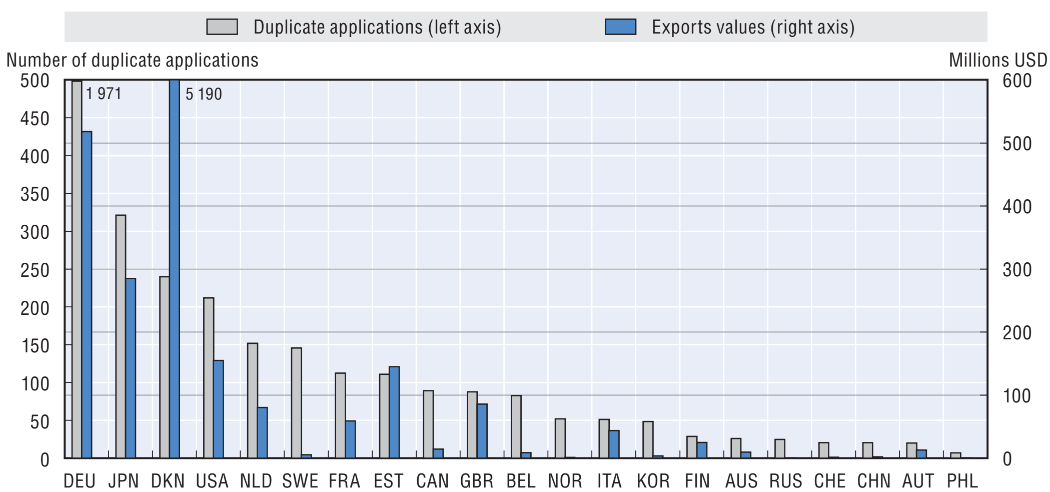
Moreover, relative to measures which rely on commodity and sectoral classifications, patent data has the great advantage that the International Patent Classification system (the IPC) is "technological" by nature. This allows for the identification of very specific "environmental" technologies – *i.e.* a distinction can be drawn between air pollution control devices designed to reduce NO_X emissions and devices designed to control SO_2 emissions (see, for example, Popp 2005). In addition, each application can list multiple codes (unlike commodity or sectoral classifications), which allows for refined searches when innovations are horizontal in nature (*i.e.* the development of fuel cells for mobile uses). And finally, unlike other data, keyword searches can be used to refine the data.

The potential to use patent data as the base from which to develop a proxy measure of technology transfer arises from the fact that protection for the invention may be sought in a number of countries.[4] While the vast majority of inventions are only patented in one country (often that of the inventor, particularly for large countries), some are patented in several countries (*i.e.* the "international patent family size" is greater than one). Such "duplicate" applications can then be used to develop indicators of technology transfer. For example, evidence on the extent of globalisation of the environmental technology sector can be developed on the basis of extractions of data on "priority" (the first patent office at which an application for a particular invention is filed) and "duplicate" (all subsequent patent offices at which protection for the same invention is sought) applications from the *PATSTAT Database*.

Of course, a patent only gives the applicant protection from potential imitators. It does not reflect actual transfer of technologies. In addition, in some cases inventions may be protected in specific markets for strategic reasons, perhaps even discouraging transfer. Despite these qualifications, there is evidence that patent applications can be used as a measure of transfer (Eaton and Kortum, 1996). Moreover, if applying for protection did not cost anything, inventors might patent widely and indiscriminately. However, patenting is costly – both in terms of the costs of preparation of the application and in terms of the administrative costs and fees associated with the approval procedure (see Helfgott, 1993 for some comparative data; Van Pottelsberghe de la Potterie and Francois, 2009) also provide more recent data for European Patent Office applications). Evidence indicates that this is true, even within the European patent system. Harhoff *et al.* (2009) find that validation and renewal fees and translation costs affect the likelihood of protecting inventions in multiple markets. If enforcement is weak, the publication of the patent in a local language can also increase vulnerability to imitation (see Eaton and Kortum, 1996; Eaton and Kortum, 1999). As such, inventors are unlikely to apply for patent protection in a second country unless they are relatively certain of the potential market for the technology that the patent covers.

Next we investigate the robustness of using duplicate patent applications to measure patterns and rates of international technology transfer. This is done through a comparison of particular areas in which trade and patent classifications are similar, including wind power and motor vehicles. Using data from the *UN COMTRADE Database* (*http://comtrade.un.org*), it is possible to compare exports of "wind-powered electric generating equipment" (HS *850231*) with the count of duplicate patent applications by priority office for "wind motors" (IPC *F03D 1-11*). Figure A.7 provides data for the main inventing countries for the period 1996-2003 – the only years for which the trade data is also available.

Figure A.7. **Number of duplicate patent applications and export of wind power technologies**



While the correlation is not perfect, it is positive and significant. Indeed the top four exporters are also the top four priority patenting offices, and the Spearman rank correlation coefficient for the top 30 countries by trade is 0.68. Some of the observed discrepancies between the two data sets may also be attributable to shortcomings in COMTRADE's coverage. For instance, for reasons of commercial confidentiality, trade figures for low-level HS classifications may be significantly downward-biased. This would explain the number of countries with no apparent exports who are known to be active in the field (for example, Sweden, Canada, Norway and Switzerland) (see *http://comtrade.un.org/kb/attachments/1.%20UN%20Comtrade%20Coverage%20and%20Limitations-GUIDbecc0aa5044f44b5a048a8b45bce6d19.pdf*).

Another area in which there is a close "marriage" between IPC and HS classifications relates to the manufacture of motor vehicles. In this case, the correlation between bilateral exports and duplicate patent applications over the period 1988-2005 is approximately 0.74 when a small number of outliers (two of 825 observations) are removed. While this area has little to do with environmental technologies *per se*, it gives a good indication of the value of duplicate patent applications as a measure of transfer.

For this corroboration exercise, road vehicle data has been chosen as it is an area with large numbers of patents, large measurable trade flows, and comparable definitions. Road vehicle patent and trade data has been used to investigate the robustness of using duplicate patent applications to measure technology transfer.

Box A.2. **Patent flow**

Patents have been extracted from the *PATSTAT Database* for road vehicles with an International Patent Classification of B62 "Land vehicles for travelling otherwise than on rail". There are three ways to measure patent flow between countries:

1. Inventor country to priority office.
2. Inventor country to duplicate office.
3. Priority office to duplicate office.

Type three, priority office to duplicate office flow counts, are used in the analysis here.

Box A.3. **Trade flow**

Data has been extracted from the *UN COMTRADE Database*. Data was extracted for the period 1988 to 2005, using both the HS and SITC classification systems, SITC = 78 "Road Vehicles" and HS = 87 "Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof". Both exports (gross exports less re-exports) and imports (gross imports less re-imports) were investigated.

Trade and patent transfers between countries are collected for the period from 1988 to 2005. Correlations between patent flows and trade over time and country pairs are shown below. There was trade data and patent transfer counts available for approximately 50 exporting countries and 60 importing countries. Results are presented using all non missing pairs and also with outliers removed, namely flows from Canada to the United States and from Japan to Germany.

Table A.1. **Correlations between trade values and counts of duplicate patent applications**

Correlation between trade flows and duplicate patenting	Full sample	Sub-sample excluding outliers (excl. CA-US and JP-DE flows)
Base dataset – all country pairs and all years (1988-2005), corr; (exports, patents)	0.47 (4 384)	0.69 (4 348)
For each country pair, aggregate over time, corr. (exports, patents)	0.52 (825)	0.74 (823)
For each exporter country, aggregate across partner countries, corr. (exports, outgoing patents)	0.76 (446)	0.87 (446)
For each importer country, aggregate across partner countries, corr. (imports, incoming patents)	0.71 (634)	0.76 (634)

Note: Pearson correlation coefficients; Number of observation in parentheses; When trade data was deflated by the US PPI all correlations improved marginally (by 0.01).

Patent transfer counts between priority office and duplicate office are extracted from the *PATSTAT Database*[5] while trade data between countries comes from the *UN COMTRADE*

Database.[6] Trade and patent transfers are collected for the period from 1988 to 2005 where there is data available for approximately 50 countries.

To substantiate the use of patent data as a measure of technological transfer, we would expect trade and patent flows to be strongly positively correlated as indeed they are found to be. Firstly, each export-import pair (1988-2005) is strongly correlated at 0.69. Aggregating over time for each export-import pair gives a correlation of 0.74. And finally, when aggregating trade and patent data for each exporter (regardless of who imports) gives a correlation of 0.88.[7]

Technological transfer occurs via many channels, though arguably trade, foreign-direct investment and licensing are the most important. Given the lack of suitable data in these areas, in particular with respect to the environmental, patent transfer data which relates to these three channels of international technology transfer, offers a suitable indicator.

Development of indicators of innovation and transfer of environmental technologies based on patent data

Patent classification systems

The International Patent Classification system (IPC, or just IC), developed at the World Intellectual Property Organisation (WIPO), is a hierarchical system classifying inventions into more than 70 000 technological groups and subgroups. It is periodically revised in order to reflect the latest technological advances. Patent offices sometimes use their own classification systems to complement the use of the IPC. For example, the European classification system (ECLA, or just EC) is an extension of the IPC with about twice as many classification codes. EPO examiners also use a further extension of the ECLA referred to as ICO codes (in-computer-only). Other classification systems include the US patent classification (USPC) or the Japanese F-terms. The work presented in this book relies on the use of IPC codes, at times complemented with ECLA codes.

Identification of environment-related technologies using patent data

Since innovation in environment-related technologies (ENV-tech) only represents one small aspect of innovation in general, prior to data retrieval from a patent database a *search strategy* must be developed that identifies the relevant patent documents using alphanumeric codes of the IPC system. For example, all patent documents with the code "*B01D 53/50 – Chemical or biological purification of waste gases; Removing sulfur oxides*" could be categorised as "SO_X end-of-pipe pollution abatement".

Development of a *search strategy* is thus based on identification of relevant patent classes that correspond to the selected "environmental" technology field. As a first step this involves an extensive review of the trade and academic literature which relates to a specific technological field. The relevant IPC classes which correspond to the different fields are then identified in two alternative ways: First, we review closely the descriptions of the classes online to find those which are appropriate (*www.wipo.int/classifications/ipc/ipc8/?lang=en*). Second, using the online world patent search engine maintained by the European Patent Office (*www.espacenet.com*), we search patent titles and abstracts for relevant keywords. The IPC classes corresponding to the patents that emerge are included, provided their description confirms their relevance.

However, in some cases, it may not be possible to identify IPC classes that alone represent the "environmental" field of interest. Possible solutions include: *i)* combining multiple "co-classes" using logical operators (that is, IPC classes whose intersection or negation yields the desired outcome); or *ii)* relevant patent classes may need to be combined with the use of keywords within the search algorithm. As part of the extraction, abstracts of patent documents with the relevant IPC classes would then be searched for the keywords identified. In some cases this would involve the exclusion of patent documents which include a particular word. However, availability of English abstracts limits the practicality of this approach when pooling data from multiple patent offices, unless the propensity to include English abstracts can be corrected for reliably.

When applying the *search strategy*, two possible types of error may arise: irrelevant patents may be included or relevant ones left out. The first error happens if an IPC class includes patents that do not bear the desired "environmental" focus. In order to avoid this problem, we carefully examine a sample of patent abstracts for every IPC class considered for inclusion, and exclude those classes that do not consist only of patents related to "environment". The second error – relevant inventions are left out – is less problematic. We can reasonably assume that all innovation in a given field behaves in a similar way and hence our extracted datasets can be seen at worst as good proxies of innovative activity in the field being considered. However, overall innovative activity may be underestimated, and the totals may be less reliable than trends.

Patent database

Over the last several years, the OECD Directorate for Science, Technology and Industry, jointly with other members of the OECD Patent Statistics Taskforce,[8] have developed a patent database that is suitable for statistical analysis – the *OECD Patent Statistics Database*. Further work has recently been undertaken by the Taskforce members towards developing a world-wide patent database – The *EPO Worldwide Patent Statistical Database (PATSTAT)*. The European Patent Office (EPO) has taken over responsibility for development and management of the database.

The *PATSTAT Database* is drawn directly from the EPO's master database (Rollinson and Lingua, 2007). It has been developed specifically for use by governmental/ intergovernmental organisations and academic institutions, and optimised for use in the statistical analysis of patent data. It has become a primary source of patent data information for statisticians, academics, and policy advisors (Rollinson and Heijnar, 2006).

The *PATSTAT Database* (EPO, 2010) has a world-wide coverage containing data from over 90 patent offices, spanning a time period stretching back to 1880 for some countries. This includes patent documents from the EPO, USPTO, JPO and other national and regional patent offices, as well as international patent applications filed under the Patent Co-operation Treaty (PCT). Overall, over 70 million patent documents are included. The database is updated on a regular basis biannually. Patent documents are categorised using the International Patent Classification (IPC) and some national classification systems (ECLA).In addition to the basic bibliometric and legal data, the database also includes patent descriptions (abstracts), applicant and inventor names, as well as citation data. The *PATSTAT Database* is thus an ideal source of patent data information for the purposes of this report.

Construction of indicators of invention in ENV-tech

Indicators of invention can be constructed as frequency counts of patent applications, disaggregated by:

- *technological field* (based on a patent search strategy, for example using IPC classes);[9]
- *priority date* (based on the first application filing date world-wide);[10]
- *inventor country* (country of residence of the inventor(s),[11] generated as fractional counts;[12])
- application authority (patent office);[13] and possibly also
- *document type* (singular, claimed priority, duplicate – based on patent family data).

Concerning the latter point, using data on patent family the following types of documents are distinguished: *singular* is patent applied for at a single office, with no subsequent applications elsewhere (*i.e.* patent family size = 1); *claimed priority* (CP) is patent application that has subsequently been claimed as priority elsewhere in the world; in other words, these are inventions that have been applied for protection in multiple countries (patent family size > 1); and finally, *duplicates* are the additional applications (sometimes referred to as *equivalents*).

There are several alternative approaches that can be used to construct the statistics:

- Count of all patent applications deposited at a single patent office (for example, at the EPO).
- Count of all patent families world-wide.
- Count of all "high-value" patent families (claimed priorities) world-wide.

In the first case, the indicator provides a count of patent application deposited at a selected patent office – including all three types of patent documents (singulars, claimed priorities, and duplicates).[14] Alternatively, only claimed priorities and duplicates could be included because, other things being equal, these should be the inventions of higher value (see discussion below). Frequently, it is patenting activity at the EPO that is studied because: *a)* the data is most complete and of best quality (within PATSTAT); and *b)* being one of the triadic offices, the statistic should reflect the "global" trends in patenting rather well.

The second indicator reflects the count of all simple patent families (that is, unique patented inventions) deposited at any office world-wide. This is achieved by counting two types of patent documents – singulars and claimed priorities. Since data from multiple patent offices are pooled together, excluding duplicates ensures that inventions are not double-counted. The upside is that this statistic is truly world-wide in its coverage because the entire stock of patent priorities is considered.

The third indicator counts "claimed priorities" (CPs) world-wide.[15] It can be argued that this statistic is the most suitable for the purpose of international comparisons because only the "high-value" priority applications are counted. The reason that claimed priorities can be viewed as representing inventions of higher value is that patenting is costly (*e.g.* translation and maintenance fees). As such, a firm will only go abroad to protect its intellectual property if it expects that the commercial value of its invention justifies it. (For empirical evidence supporting this argument see Guellec and van Pottelsberghe, 2000; Harhoff *et al.*, 2003. Indeed, the use of an indicator which excludes the "one-member" patent families – that is, an indicator based on CPs – justified on the grounds of an "economic threshold criterion" was advocated already by Faust and Schedl 1983, and Faust, 1990.) Moreover, by excluding priority applications which have never been claimed abroad (one-member families, or singulars) this approach may help contain concerns over strategic patenting. In addition (Faust, 1990), it will thus exclude the large number of exclusively domestic Japanese patent applications with usually only one claim.[16]

It must be noted that the OECD triadic patent family (TPF) indicator has too, been developed with the specific purpose to allow international comparisons. However, in the context of a narrowly-defined technological field – such as environment-related technologies – the TPF indicator is overly restrictive leaving little variation in the data. The less-restrictive CPs thus provide a good alternative.

Care needs to be taken when conducting descriptive and econometric analyses. In particular, comparisons across inventor countries should take into account the potential "home bias" of domestic inventors. In addition, the propensity of inventors to patent, the breadth of invention claims covered by a patent, as well as the scope of patent protection, each vary over time and across countries. In order to account for these differences, the patent counts representing "environmental" innovations should therefore be expressed as shares, or these factors should be controlled for econometrically, using data on patenting activity overall.

Construction of indicators of international transfer of ENV-tech

Indicators of technology transfer can be constructed as frequency counts of:

1. **Patent families, disaggregated by:**
   - *technological field* (based on a patent search strategy);
   - *time* (based on priority date or application date);
   - *source country* (based on inventor country); and
   - *recipient country* (based on application authority – including both the priority and duplication offices).
2. **Duplicate patent applications, disaggregated by:**
   - *technological field* (based on a patent search strategy);
   - *technological field* (based on a patent search strategy);
   - time (based on priority date or application date);
   - source country (based on priority office); and
   - recipient country (based on duplication office).

In sum, in the former case the statistic measures transfer between inventor country and the patent offices included in the corresponding patent family.[17] In the latter case, transfer is to occur from (the country of) the priority office to the duplication office.[18]

Notes

1. This work would not have been possible without collaboration with the Economic Analysis and Statistics Division of the OECD Directorate for Science, Technology and Industry. In addition, much of the groundwork on developing the indicators has benefited from collaboration with researchers at universities and research institutes elsewhere. "Acknowledgement" is provided in the relevant chapters.
2. The data is available for download at *http://stats.oecd.org/Index.aspx?DatasetCode=PATS_IPC*. For metadata and future updates see *www.oecd.org/environment/innovation/indicator*.
3. The distribution of R&D expenditures is set out in the OECD *Frascati Manual*: "The Measurement of Scientific and Technological Activities: Proposed Standard Practice for Surveys on Research and Experimental Development" (OECD, 2002). The definition of the socio-economic objective (SEO) for "control and care for the environment" covers research into the control of pollution, aimed at the identification and analysis of the sources of pollution and their causes, and all pollutants, including their dispersal in the environment and the effects on man, species (fauna, flora, micro-organisms) and the biosphere. Development of monitoring facilities for the measurement of all kinds of pollution is also included. The same is valid for "the elimination and prevention of all forms of pollution in all types of environment".
4. See Dernis *et al.* (2001). Note that the information provided in patent documents can also be used to measure both market (co-invention, co-ownership) and non-market (citation) international knowledge flows. See Guellec and van Pottelsberghe de la Potterie (2000) for a discussion and evidence derived from patent data on the internationalisation of knowledge and technology flows.
5. IPC Code of B62 – Land vehicles for travelling otherwise than on rail.
6. Using both the HS and SITC classification systems, SITC = 78 "Road vehicles" and HS = 87 "Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof".
7. Results are presented using all non missing pairs and also with outliers removed, namely flows from Canada to the United States and from Japan to Germany.
8. Other Taskforce members include the European Patent Office (EPO), the Japan Patent Office (JPO), the United States Patent and Trademark Office (USPTO), the World Intellectual Property Organisation (WIPO), the US National Science Foundation (NSF), Eurostat, and the European Commission Directorate-General for Research.
9. For a list of IPC codes and their definitions see *www.wipo.int/classifications/ipc/ipc8*.
10. "Priority date" indicates the earliest application date worldwide (within a given patent family).
11. For a list of two-letter country codes see *www.wipo.int/standards/en/pdf/03-03-01.pdf*.
12. Generating the counts as "fractional" means that if inventors from two (three, or more) different countries are involved, only a fraction of 0.5 (0.33, etc.) will be counted for a given patent application.
13. See *www.wipo.int/standards/en/pdf/03-03-01.pdf* for a list of two-letter codes of application authorities.
14. For example, this approach was adopted in Johnstone *et al.*, 2010.
15. This is the approach adopted in Chapters 2 and 3.
16. CPs account for a relatively small proportion of the stock of patent applications. For example, in a study focusing on innovation in climate change mitigation technologies Haščič *et al.* (2010) find that only about 11% of the relevant stock of patent applications included in PATSTAT were CPs, with 34% being their duplicates, and 55% being singulars. In other words, in climate change mitigation technologies CPs represent only 16% of the stock of patented inventions (simple patent families), while the large majority (84%) of these inventions were only protected at a single patent office (singulars). It must be noted that there is variation in these proportions across patent offices.
17. For example, this approach was adopted by Dechezleprêtre *et al.* (2011).
18. This is the approach adopted in Chapter 4.

References

Dechezleprêtre, A., M. Glachant, I. Haščič, N. Johnstone and Y. Ménière (2011), "Invention and Transfer of Climate Change Mitigation Technologies: A Global Analysis", *Review of Environmental Economics and Policy* (forthcoming 2011).

Dernis, H., D. Guellec and B. van Pottelsberghe de la Potterie (2001), "Using Patent Counts for Cross-country Comparisons of Technology Output", *STI Review*, No. 27, pp. 129-146, OECD.

Eaton, J. and S. Kortum (1996), "Trade in Ideas: Patenting and Productivity in the OECD", *Journal of International Economics*, No. 40, pp. 251-278.

Eaton, J. and S. Kortum (1999), "International Technology Diffusion: Theory and Measurement", *International Economic Review*, Vol. 40(3), pp. 537-570.

European Patent Office (EPO) (2010), *EPO Worldwide Patent Statistical Database (PATSTAT)*, April 2010 edition, European Patent Office.

Faust, K. (1990), "Early identIfication of Technological Advances on the Basis of Patent Data", *Scientometrics*, Vol. 19(5-6), pp. 473-480.

Faust, K. and H. Schedl (1983), "International Patent Data: Their Utilisation for the Analysis of Technological Developments", *World Patent Information*, Vol. 5(3), pp. 144-157.

Guellec, D. and B. van Pottelsberghe de la Potterie (2000), "Applications, Grants and the Value of a Patent", *Economics Letters*, No. 69, pp. 109-114.

Griliches, Z. (1990), "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature*, Vol. 28, No. 4, pp. 1661-1707.

Harhoff, D., K. Hoisl, B. Reichl and B. van Pottelsberghe de la Potterie (2009). "Patent Validation at the Country Level – The Role of Fees and Translation Costs", *Research Policy*, Vol. 38(9), pp. 1423-1437.

Harhoff, D., F.M. Scherer and K. Vopel (2003), "Citations, Family Size, Opposition and the Value of Patent Rights", *Research Policy*, No. 32, pp. 1343-63.

Haščič, I., N. Johnstone, F. Watson and C. Kaminker (2010), "Climate Policy and Technological Innovation and Transfer: An Overview of Trends and Recent Empirical Results", *OECD Environment Working Paper*, No. 30, *www.oecd.org/env/workingpapers*.

Helfgott, S. (1993), "Patent Filing Costs Around the World", *Journal of the Patent and Trademark Office Society*, July, pp. 567-580.

Johnstone, N. (ed.) (2007), *Environmental Policy and Corporate Behaviour*, Edward Elgar.

Johnstone, N., I. Haščič and D. Popp (2010), "Renewable Energy Policies and Technological Innovation: Evidence Based on Patent Counts", *Environmental and Resource Economics*, Vol. 45(1), pp. 133-155.

Maskus, K.E. (2004), "Encouraging International Technology Transfer", *ICSTD/UNCTAD Issue Paper*, No. 7 of UNCTAD-ICSTD Project on IPRs and Sustainable Development.

Meyer, M. (2002), "Tracing Knowledge Flows in Innovation Systems", *Scientometrics*, Vol. 54, No. 2, pp. 193-212.

OECD (2009), *OECD Patent Statistics Manual*, OECD, Paris.

OECD (2008a), *Main Science and Technology Indicators*, OECD, Paris.

OECD (2008b), *Environmental Policy, Technological Innovation and Patents*, OECD, Paris.

OECD (2007), *Issues of Dual Use and Reviewing Product Coverage of Environmental Goods*, OECD, Paris.

OECD (2002), "The Measurement of Scientific and Technological Activities: Proposed Standard Practice for Surveys on Research and Experimental Development", *Frascati Manual*, OECD, Paris.

OECD (2001), *Environmental Goods and Services: The Benefits of Further Global Trade Liberalization*, OECD, Paris.

OECD (1999), *The Environmental Goods and Services Industry: Manual for Data Collection and Analysis*, OECD, Paris.

Popp, D. (2005), *Using the Triadic Patent Family Database to Study Environmental Innovation*, [ENV/EPOC/WPNEP/RD(2005)2], OECD, Paris.

Rollinson, J. and R. Heijna (2006), *EPO Worldwide Patent Statistical Database (PATSTAT)*, presentation at the EPO and OECD Conference on Patent Statistics for Policy Decision Making, Vienna, Austria, October.

Rollinson, J. and D. Lingua (2007), "The Development of PATSTAT", presentation at the EPO and OECD Conference Patent Statistics for Policy Decision Making, Venice, Italy, October.

Sinclair-Desgagne, B. (2008), "The Environmental Goods and Services Industry", in *International Review of Environmental and Resource Economics*, Vol. 2, Issue 1, pp. 69-99.

Steenblik, R. (2003), "Environmental Goods: A Comparison of the OECD and APEC Lists", *OECD JWTE Working Paper*, No. 2005-04.

van Pottelsberghe de la Potterie, B. and D. Francois (2009), "The Cost Factor in Patent Systems", *Journal of Industry, Competition and Trade*, No. 9, pp. 329-355.