OECD Science, Technology and Industry Working Papers
2015/04

Identifying and inducing breakthrough inventions: An application related to climate change mitigation

Florian Egli,

Nick Johnstone,

Carlo Menon

https://dx.doi.org/10.1787/5js03zd40n37-en

OECD

OECD SCIENCE, TECHNOLOGY AND INDUSTRY WORKING PAPERS

The release of this working paper has been authorised by Andrew Wyckoff, OECD Director for Science, Technology and Innovation.

IDENTIFYING AND INDUCING BREAKTHROUGH INVENTIONS: AN APPLICATION RELATED TO CLIMATE CHANGE MITIGATION

Florian Egli[i], Nick Johnstone[ii], and Carlo Menon[ii]

OECD Paris

ABSTRACT

Most of the projections of the cost of meeting climate change mitigation targets hinge crucially upon assumptions made about the cost and timing of the development of breakthrough technologies. However, very little is known about the conditions which are likely to give rise to breakthrough technologies. This paper seeks to uncover attributes of inventions – as reflected in patent data – which serve as "leading indicators" of subsequent technological and market development in climate change mitigation technologies. The role of industrial generality emerges as being robustly correlated with subsequent technological diffusion, whether measured as subsequent patent counts, commercial applicability, or attractiveness to risk finance. The indicator of closeness to science shows also a positive association with later technological diffusion. Originality and radicalness have more ambiguous results. This work can be seen as a foundation for the future development of a methodology providing guidance to policymakers in the choices made with respect to public support for different technological fields.

Keywords: Climate change mitigation technologies; Green growth; Breakthrough inventions; patent quality indicators.

JEL Classifications: O31, O33, Q54, Q55

EXECUTIVE SUMMARY

In the past, a number of technologies (e.g. information and communication technologies and nanotechnologies) have had far-reaching consequences on important economic variables (e.g. productivity) and/or our capacity to meet social and environmental challenges at relatively lower cost. The early identification of technologies which are likely to have far-reaching economic consequences would be of considerable value to policymakers.

Drawing upon the more general literature on the "quality attributes" of inventions, this paper seeks to uncover some of the essential features of breakthrough technologies in a specific domain. The definition of breakthrough technologies adopted is rather wide and include new, fast-growing, radical technologies that either introduce new product or process with very high market potential, or that make existing established technologies rapidly obsolete. While the focus of the paper is on climate mitigation technologies, the analysis is relevant for other technology fields as well, and this paper should be seen as a case study application. The intention is to try and uncover attributes of inventions (as reflected in patent data), which serve as "leading indicators" of subsequent technological and market development. This work can thus be seen as a foundation for the future development of a methodology, which can provide guidance to policymakers in the choices made with respect to public support for different technological fields.

The choice of climate change mitigation technologies is motivated by a number of factors. Firstly, if the international community is to meet the declared objectives under the Intergovernmental Panel on Climate Change (IPCC) – global warming limited to 2^{O} C above pre-industrial levels – radical innovations will be required. Secondly, most of the projections of the cost of meeting such targets hinge crucially upon assumptions made about the cost and timing of the development of breakthrough technologies. And finally, patent examiners at the European Patent Office have recently developed search algorithms to identify relevant technologies, many of which are potential breakthrough technologies.

Unfortunately, very little is known about the conditions which are likely to give rise to breakthrough technologies. This paper builds on a sub-set of the attributes proposed and discussed in Squicciarini, Criscuolo and Dernis (2013)[iii] which are consistent with the notion of an invention which is ground-breaking and which then subsequently is widely diffused. More particularly the attributes used for the analysis are:

- **Originality** – an indication of the "breadth" of the technology fields on which a patent relies.
- **Radicalness** – an indication of the extent to which a patent relies on previous inventions from fields other than its own.
- **Industrial Generality** – an indication of the range of sectors of firms who subsequently cite a given patent.
- **Family size** – an indication of the number of markets in which a patent is protected.
- **Closeness to Science** – an indication of the extent to which a patent draws upon the scientific literature rather than patents.

The analysis finds that *photovoltaic energy generation* ranks high in terms of industrial generality and closeness to science, but low in terms of radicalness. *Hydrogen technology* emerges as another interesting case, ranking highly in terms of industrial generality, and relatively highly in all of the other measures. *Biofuels* rank highly for all indicators as well, and in particular with respect to closeness to science and family size. Interestingly, some of the most widespread technologies – and widely subsidised – such as *hydro energy*, *wind energy* or *nuclear fission* show lower values than the counterfactual average technology for all the indicators.

In preliminary empirical analysis, the role of industrial generality emerges as being particularly important, with positive implications for downstream success whether measured as subsequent patent counts, commercial applicability, or attractiveness to risk finance. The indicator of closeness to science has positive and significant effects, except in the case of a measure of applicability based on the share of patents applied for by the private sector. Originality and radicalness have more ambiguous results, with signs changing depending upon the measure of success. In summary, the only robust indicator of success for technologies is industrial generality. In every specification a more “general” technology predicts a growth in patent applications, granted patents, relative business propensity to patent, and risk finance attraction.

It should be emphasised that providing discretionary support is a hazardous exercise. There is little doubt that the backbone of innovation policy should be technology-neutral. However in practice many governments are providing discretionary incentives for specific technologies, and this report seeks to explore in a very preliminary manner how such choices can be made in a more evidence-based manner. Even if further work will allow for the identification of reliable leading indicators of future technology and market developments, there will still be great uncertainty. As such, it is important to bear in mind the importance of ‘trial-and-error’ and the establishment of transparent and unambiguous exit mechanisms. In fact, the government needs to exit not only if a technology proves to be unsuccessful, but also if a technology proves to be successful enough for its future developments to be driven by private actors.

TABLE OF CONTENTS

IDENTIFYING AND INDUCING BREAKTHROUGH INVENTIONS: AN APPLICATION RELATED TO CLIMATE CHANGE MITIGATION[iv]

1. Introduction

In the past, a number of technologies have had far-reaching consequences on important economic variables (e.g. productivity) and/or our capacity to meet social and environmental challenges at relatively lower cost. In recent years the case of information and communication technologies (ICT) is one such example. Important advances in medical sciences related to nanotechnology are another example. The early identification of technologies which are likely to have far-reaching economic consequences would be of considerable value to policymakers.

This paper explores these issues in a preliminary manner, building on work undertaken in the OECD Directorate for Science, Technology and Innovation on the qualitative characteristics of inventions based on information contained in patent applications (see Squicciarini, Dernis and Criscuolo, 2013; henceforth SDC). More specifically, drawing upon the more general literature on the “quality attributes” of inventions, this paper seeks to uncover some of the essential features of breakthrough technologies in the environmental domain. The results of the econometric analysis suggest that “industrial generality”, i.e. an original measure of the sectoral breadth of the utilisation of patents by private companies, is strongly and robustly associated with later rapid diffusion of a given technology. Other characteristics of the average patents – including originality, radicalness, and closeness to science – provide more varied results and do not show a clear-cut robust association with later diffusion patterns.

In this paper, the working definition of breakthrough technologies adopted is rather wide. It includes new, fast-growing, radical technologies that either introduce new process or product innovations with high market potential, or that make existing established technologies rapidly obsolete. In the climate change mitigation literature such technologies are usually referred to as backstop technologies, defined as a “new technology producing a close substitute to an exhaustible resource by using relatively abundant production inputs and rendering the reserves of the exhaustible resource obsolete when the average cost of production of the close substitute falls below the spot price of the exhaustible resource” (Dasgupta and Heal, 1974).

While the focus of the paper is on climate mitigation technologies, the analysis is relevant for other technology fields as well, and this paper should be seen as a case study application. The intention is to uncover attributes of inventions (as reflected in patent data) which serve as “leading indicators” of subsequent technological and market development. The objective is to devise a methodology which can provide guidance to policymakers in the choices made with respect to public support for different technological fields. This is a hazardous exercise, and there is little doubt that the backbone of innovation policy should be technology-neutral (i.e. protection of intellectual property, investment in basic research, fostering collaboration between different actors in the innovation process, etc.). However, as a complement to this general and technology-neutral policy framework, in practice many governments are providing discretionary incentives for specific technologies, and this report seeks to explore in a very preliminary manner how such choices can be made in a more evidence-based manner.

For many environmental concerns continuous incremental innovations are likely to be the most economically efficient technological trajectory. However, for at least some environmental concerns, meeting stated environmental objectives is likely to require more radical breakthrough innovation. The case of climate change stands out as an example. Due to the long-lived nature of greenhouse gases (GHGs) as stock pollutants, concentrations are likely to continue to rise for many decades, even with the realisation of significant emission reductions in short- to medium-run. Therefore, if the international community is to meet the declared objectives under the Intergovernmental Panel on Climate Change (IPCC) – global warming limited to 2^{O} C above pre-industrial levels – more radical innovations will be required.

Moreover, most of the projections of the cost of meeting such targets hinge crucially upon assumptions made about the cost and timing of the development of breakthrough (or backstop) technologies. Unfortunately, very little is known about the conditions which are likely to give rise to breakthrough technologies. This is hardly surprising in light of the fact that our knowledge of the determinants of innovation more generally remains imperfect. Forecasting the emergence of breakthrough technologies in decades to come and modelling their economic implications is correspondingly more difficult. The combination of heightened technological, economic, and policy uncertainty compounds the difficulties of the exercise for the researcher. Indeed, in the computable general equilibrium (and other) models which seek to forecast the costs of meeting climate change objectives the emergence of backstop/breakthrough technologies is modelled in an *ad hoc* manner (See Löschel and Schymura 2013 for a review).

The paper is structured as follows. Section II provides the policy research motivation for the paper. Section III presents the conceptual basis for the indicators and provides technical information on the methodology used for their construction. Section IV provides an overview of descriptive evidence related to the different technology fields covered in the report. The model and the results are then provided in Section V, followed by the conclusions.

2. The Motivation for the Analysis of Breakthrough Environmental Technologies.

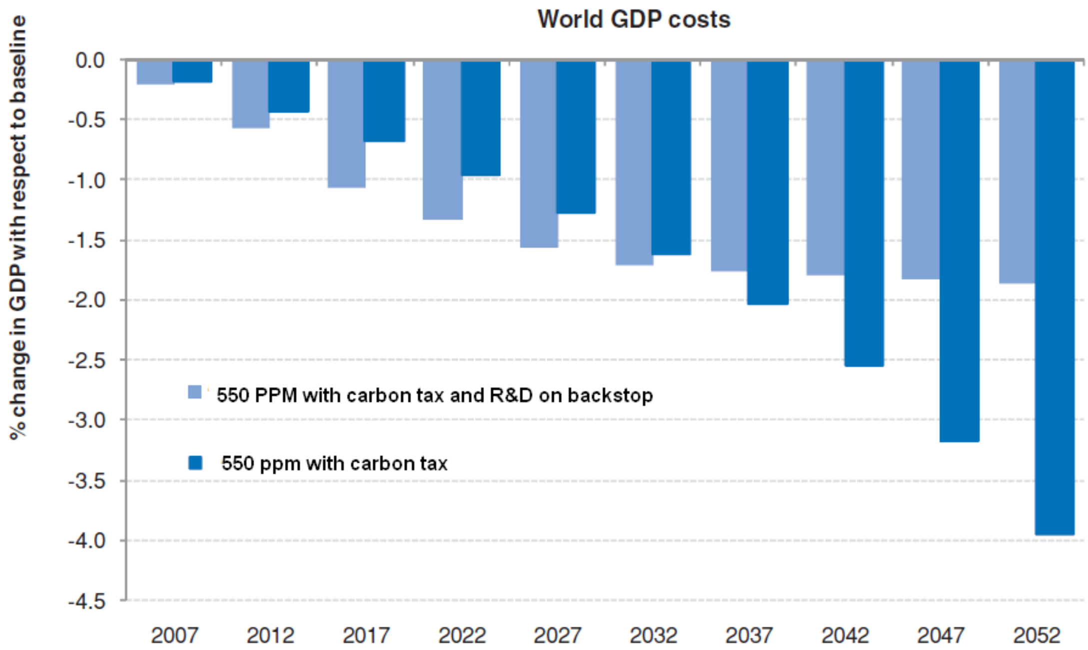
The potential importance of "breakthrough" environmental technologies is reflected in much of the work that has been undertaken in recent years on the costs of meeting GHG mitigation objectives. Relative to other policy domains, assessing the costs and benefits of climate change policy is complicated by the joint existence of a number of factors, including:

- Non-market nature of many of the benefits;
- Long-run nature of both benefits and costs of policy interventions and the difficulties of political commitment in the long term; and,
- Presence of irreversibilities (both economic and ecological)[v] and fundamental uncertainty.

This paper is only concerned with the latter point, and specifically with uncertainty about the costs of meeting stated/agreed environmental objectives over the longer-term. And with respect to this issue a key question is the forecasting of the costs and quality of emerging breakthrough mitigation technologies.

Indeed, it is these assumptions which "drive" many of the results in the literature. For example, Bosetti et al. (2011) on the benefits of targeted R&D on "far-from-market" technologies which are under research but not yet viable (e.g. advanced biofuels, nuclear and fuel cells). In one scenario the general equilibrium model (WITCH) is run in which the entire burden of meeting the objective of stabilisation at 500 PPM is met with a carbon tax. In another scenario, R&D expenditures are devoted to these breakthrough technologies with assumed impacts on the cost of abatement. The results (see Figure 1) reveal the importance of the latter assumptions. While costs are higher in the short-run, in due course the costs are approximately halved.

Figure 1. Percentage change in GDP with and without R&D on "Backstop" Technologies



Note: The graph reports two alternative scenarios on the cost of climate-change mitigation, respectively with or without backstop technologies.

Source: OECD Economics of Climate Change Mitigation. Based on research undertaken by V. Bosetti et al. (2010) in http://www.feem.it/userfiles/attach/2010471754234NDL2010-042.pdf.

In effect, assumptions concerning backstop technologies "drive" existing estimates of the costs of climate mitigation beyond 2050. Moreover, there is good reason to believe that governments are likely to under-invest in such technologies relative to other domains. Part of this is attributable to the absence of socially optimal levels of carbon taxation. However, even in the presence of such levels of taxation there is likely to be under-investment in the kinds of technologies which bring about far-reaching structural transformation of the economy for at least three reasons:

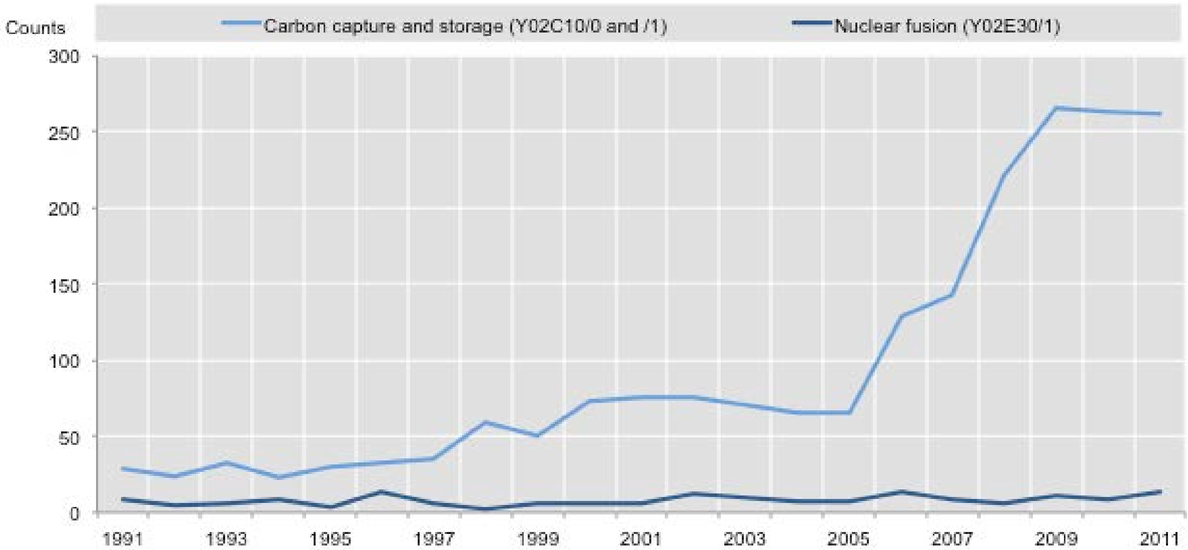
i) "general-purpose technologies" are likely to have particularly high levels of positive spillovers on other technologies, inducing innovation across the economy;

ii) "radical" innovations for which there is a tail of particularly low probability but potentially very high return investments are notoriously under-financed; and

iii) "public good technologies" in which public policy plays an important role in shaping the market, are subject to a level of regulatory risk which compounds commercial risk.

The key question – in a context of time and public budget constraints - is then the identification of "which" technologies. As noted above in their modelling of the costs of climate mitigation Bosetti et al. (2011) make assumptions about the technological trajectories of advanced biofuels, nuclear fusion and fuel cells. In the past many general equilibrium and integrated assessment models have considered carbon capture and storage (CCS).[vi] However, if policymakers were to have "bet" on (for example) either nuclear fusion or CCS just prior to the Kyoto Protocol, they would have had little guidance on which of these were more promising in terms of future knowledge development – and thus likely costs (see Figure 2).

Figure 2. Counts of Inventions for Two Common "Backstop" Technologies



Note: Counts of All Patent Priorities (Claimed Priorities and Singulars).

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) May 2013.

The challenge for the policymaker is, therefore, to be able to identify technologies which in the future have the potential to realise environmental objectives at minimum cost. There are, of course, hazards associated with providing discretionary support for particular technologies. As such, it is important to emphasise that more general policies which target the environmental externality (e.g. a carbon tax) and the knowledge externality (e.g. sound IP regimes) in a neutral way should form the core of any policy regime. Nonetheless, if space is accorded to more "directed" support it is important to identify an informationally-parsimonious methodology to identify potential candidates in a sufficiently timely manner.

3. Breakthrough Innovations and Patent Quality

3.1. Principal hypotheses

Much of the recent work on technological innovation in the context of mitigation of climate change draws upon patent data. See OECD (2012) and Johnstone and Haščič (2013) for recent reviews. It must be noted that all measures of innovation are imperfect. For instance, R&D data are unsatisfactory insofar as they measure an input to innovation, rather than an output. Data on scientific personnel suffers from a similar shortcoming. In recent years, bibliometric data has been used, but it can be difficult to develop efficient topical search strategies; furthermore, the link between publication and value is likely to be even more approximate than with patent data.[vii]

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, OECD 2009). Patent data provide a wealth of information on the nature of the invention and the applicant, the data are readily available (if not always in a convenient format), and discrete (and thus easily subject to the development of indicators). Most importantly, the application-based nature of the patent classification systems allows for a richer characterisation of relevant technologies. However, one needs also to keep in mind the well-known caveat that not all inventions are patented, although there are few examples of economically significant inventions that are not patented (OECD, 2001).

However, for the most part the patent data that has been applied in empirical work has been treated in a relatively undifferentiated manner. In some cases minimum thresholds on "family size" have been used as a means to remove low-value patents from the counts. For example, the papers included in OECD (2012) use the principle of claimed priorities, which is equivalent to a family size of two. Even more binding, the concept of a triadic patent family (TPF) has been developed to define family of patents for which protection has sought at the EPO, JPO and USPTO. This constrains the counts only to a subset of high-value patents, but is unlikely to be appropriate for the development of indicators for specific fields (see Haščič et al. 2015).

A small number of studies have used patent data to forecast the future development of different technologies. For instance, Daim et al. (2006) use both patent (and bibliometric) data to forecast the future development of three technologies: fuel cells, food safety and optical storage technologies. In addition, Haupt et al (2007) use a number of indicators (e.g. citations from patent and non-patent literature, duration of the patent examination process) to forecast transitions in the pattern of development of patent applications in a specific technology (pacemakers).

There is, however, much richer information that can be gleaned from patent data and recent work at the OECD has sought to develop sophisticated measures of patent quality using data extracted from PATSTAT (see SDC). Indeed, the quality attributes of a patent can take on a host of distinct meanings. These different meanings are not mutually exclusive, and different notions of quality may be relevant for different empirical questions. From an economist's perspective a "good" patent is generally one that fulfils the key objectives of the patent system, i.e. to reward and incentivise innovation while enabling diffusion and further technological developments (see Guellec and van Pottelsberghe de la Potterie, 2007). However, it is important to make clear that in this report "quality" has a completely neutral connotation and it relates to a set of characteristics of a patent. This means that the analysis does not build on assumptions that, e.g. more radical or original patents are more valuable.[viii]

More specifically, this paper focuses on a sub-set of the attributes proposed and discussed in SDC[ix] which seem to match most closely with the concept of "backstop technologies", i.e. an invention which is ground-breaking and which then subsequently is widely diffused. More particularly the attributes used for the analysis are:

- **Originality** – (Tratjenberg et al., 1997) – an indication of the "breadth" of the technology fields on which a patent relies.
- **Radicalness** – (Shane, 2001) – an indication of the extent to which a patent relies on previous inventions from fields other than its own.
- **Industrial Generality** – (modified from Bresnahan and Tratjenberg, 1995) – an indication of the range of sectors of firms who subsequently cite a given patent.
- **Family size** – (Lanjouw et al., 1998; Guellec and van Pottelsberghe de la Potterie, 2000) – an indication of the number of markets in which a patent is protected.
- **Closeness to Science** – (Tratjenberg, Henderson and Jaffe, 1997) **–** an indication of the extent to which a patent draws upon the scientific literature rather than patents.

All of these measures are constructed at the level of an individual patent (based on micro-data on patent applications) and then aggregated at the level of a patent class (8-digit and 9-digit). The extractions are based on the new Y-tags developed by patent examiners at the EPO in consultation with a number or external experts, including researchers at the OECD. The benefit of the new Y02 scheme is twofold. First, it allows for the identification of patent documents in technological fields that would otherwise not be possible by a non-specialist. Second, it provides a much greater level of detail in the classification of environmental technologies than what would be possible using the IPC. This is a significant advantage over previous searches based on the IPC. See Veefkind et al. (2012) for further details. Annex A contains the full list of classes.

The objective of this paper is to assess the extent to which different “quality” attributes affect the longer-term implications of different classes of inventions. In the context of this paper the ideal measure would be information on either the marginal cost of abatement associated with alternative technologies and/or their market penetration rates.

Unfortunately such data is not available. Three alternative variables as dependent variables are therefore used (see Section 3.3 for a detailed description):

- Growth in counts of patents within the class;
- Commercial applicability of the patents within the class – measured in terms of percentage of assignees from profit-seeking firms;
- Financial attractiveness of patents within the class – measured in terms of venture capital flows to firms with patents within the class.

Our central hypotheses are:

- H1a: More original patents are more likely to be breakthrough technologies since they draw upon a wide range of knowledge in order to address a given challenge.
- H1b: More radical patents are more likely to be breakthrough technologies since they are less likely to lead to incremental changes in which a given challenge is addressed.
- H2: More general patents are more likely to be breakthrough technologies since they are more likely to be diffused across the economy.
- H3: Patents with larger family sizes are more likely to be breakthrough technologies since they are more likely to be diffused internationally.
- H4: Patents which are closer to science (more “basic” in nature) are more likely to be breakthrough technologies since they are more likely to represent fundamental advances in knowledge. [x]

The empirical analysis relies on five quality indicators, four of which are developed in SDC, while the remaining one has been developed ad-hoc for this paper. Because of data availability, all the patent variables used in this paper are built using exclusively patents applied for at the European Patent Office. No other restrictions – e.g. in terms of applicants’ or inventors’ country – are applied.

3.2. Construction of the patent quality indicators

In detail, the four patent quality indicators mentioned above are calculated by SDC as follows:

- Radicalness (à la Shane): it measures the degree to which a patent relies (via backward citations) on a diversified array of technologies. The index is calculated as follows:

$$Radicalness_p = \sum\nolimits_j^{n_p} CT_j/n_p \,;\; IPC_{pj} \neq IPC_p$$

where CT_j denotes the count of IPC-4 digit codes (IPC_{pj}) of patent j cited in patent p, excluding those which patent p is also allocated to; while the denominator n is the count of all IPC classes (defined at the most disaggregated level) listed in the backward citations of patent p. The higher the ratio, the more diversified the array of technologies on which the patent relies upon.

- Originality: the indicator is similar to the radicalness one, and it measures the degree to which citations to different technological classes are dispersed. The index is calculated as follows:

$$Originality_p = 1 - \sum\nolimits_j^{n_p} S_{pj}{}^2$$

where s_{pj} is the percentage of citations made by patent p to patent class j out of the n_p IPC 4-digit (or 7-digit) patent codes contained in the patents cited by patent p.

- Closeness to science: calculated as the ratio of citations to non-patent literature (e.g. academic articles) in all backward citations.
- Family size: proxied by the number of patent offices at which a given invention has been protected.

The industrial generality index – developed for this paper – aims to measure the breadth of the commercial applicability of a technological class in the industrial space. The underlying idea is that a general technological class has a wider set of subsequent applications. It closely recalls the patent generality index based on forward citations developed by SDC, which in turn builds on the original idea of Trajtenberg et al. (1997).[xi]

The index is calculated using the HAN database developed at the OECD. The HAN database results from a sophisticated matching – based on string similarity – of the PATSTAT database with the commercial database ORBIS® maintained by the Bureau van Dijk, which is useful here as it also contains information on firms' main sector of activity. The index also benefits from additional refinements of the ORBIS dataset by Gal (2013) and Gonnard and Ragoussis (2013), and of additional work on the HAN database by Andrews, Criscuolo and Menon (2014).[xii]

Contrary to the four SDC indices previously discussed, the industrial generality indicator is not initially calculated at the patent level, but directly at the technological class level. For each technological class, the index corresponds to the Hirschman-Herfindahl Index (HHI) measuring the distribution of patent shares across the 3-digit NACE industry sectors of the patent applicant:

$$IGI_k = 1 - \sum\nolimits_q \left(\frac{applications_q}{applications_k}\right)^2$$

where k indexes the patent technological class and q the applicant NACE 3-digit sector. If all the patents in a given technological class were concentrated in a single 3-digit NACE industry, the index would be equal to zero. On the contrary, if the patents are assigned to a large number of industries with similar shares, the index would take a value close to one.

As compared to the traditional patent generality index, the industrial generality index has some distinctive features which are particularly advantageous in the present setting. First, it conveys information on the actual industrial exploitation of patents, making it more economically relevant with respect to indices which are based solely on patent data. Second, it is not based on future forward citation patterns, which would be an "endogenous" measure in this analysis, as technological fields that are becoming more successful over time are also expected to receive more citations to earlier patent cohorts. Third, it is less prone to a potential bias originating from the increase in the number of IPC technological classes than other indices. For instance, the originality index is by construction positively correlated with the number of different 4-digit IPC classes referring to the same Y-tag CPC classification. This may introduce spurious correlation between bursts and the originality index because when a technological field is expanding, patent examiners tend to introduce new IPC classes in order to better characterise the new developments of the technology. However, this does not affect the industrial generality index, as it is based on the NACE 3-digit classification which is fixed over time and unaffected by patent technological classification.

Table 1. Summary of Quality indicators

Indicator	Description
Industrial generality	An indication of the range of industrial sectors in which a patent is subsequently cited. A measure of width of the business field of the company applying for the patent.
Originality	An indication of the "breadth" of the technology fields on which a patent relies (*ex post*).
Closeness to science	An indication of how close an invention / technology is to academia and science as opposed to former patents and commercial developments.
Radicalness	An indication of the extent to which a patent relies on previous inventions from fields other than its own. A radical patent develops something *new in its field.*
Family size	An indication of the number of markets (hence patent offices) in which a patent is protected.

Source: elaboration from SDC.

3.3. Measures of technological classes growth and commercial diffusion

As noted above, our (imperfect) proxies for the longer-term impact of inventions in different fields relate to growth patterns in counts and the commercial diffusion of patent technological classes – i.e., the dependent variables in the econometric analysis – are measured in several different ways.

- **Patent growth:** the first two variables are based on a simple patent application count, expressed as i) the *total number of applications over the period of interest* (2001-2010 in this case); and the ii) *average annual growth rate*. Two additional indicators are calculated exactly in the same way, but replacing applications with granted patents. Similarly, patent count by technological class can be calculated as a simple or a fractional count, respectively. The fractional count takes into consideration the fact that patents can be tagged with N≥1 technological classes (without any order of priority); to account for that, each technology class tag is weighted by 1/N, making the patent growth variables reflect the actual number of patents (rather than the prevalence of a technology class in the case of single counts) and avoiding double-counting when aggregating the figures across all technological classes.

- **Commercial viability:** the following set of indicators takes into account changes in the commercial viability of patents belonging to a given technological class, and it exploits information on the institutional nature (e.g., university, government, private company, etc.) of the patent applicants. Specifically, the indicators are calculated as the share of patents in each technological class applied for by a private company, based on the classification developed by ECOOM in partnership with Sogeti (EuroStat, 2011).[xiii] Again, the indicator can be calculated for both all applications or granted patents only, and as a simple or fractional count.

- **Risk finance attractiveness:** this indicator is based on previous OECD work drawing upon a commercial database on risk finance investments in the green sector maintained by the CleanTech Group® (Criscuolo and Menon, 2014).[xiv] The CleanTech database collects information on risk-finance deals (including venture capitalist, business angel, and private equity investments) in the green sector worldwide, starting from early 2000's. The approximate investment amount is reported, together with the name and the location of the funded company. The latter pieces of information are also available in PATSTAT for patent applicants. This allows for the matching of the two datasets using string-similarity algorithms in a manner analogous to that discussed above in the construction of the industrial generality index – and retrieving the patent history of the funded companies. This in turn allows for the calculation of the total amount of investments for each technological class and year. Specifically, each patent matched to a funded company is attributed the funding that the owning company receives up to five years after the patent application date. Robustness tests are also done with alternative versions of the index based on i) a three, rather than five, year period, and ii) dividing the funding received by the total number of patents received in the same time window.

It is worth mentioning that the measure is approximate – although the CleanTech database offers a good coverage of investments, especially for later years in our sample, the matching with patent data is done through text similarity rather than through an exact numeric identifier, which inevitably leads to "false positive" (type A) and "false negative" (type B) errors, due to different name and city spelling. In addition, the implicit assumption that the company patent portfolio plays a determinant role in attracting funding from risk finance might be debatable. However, the index can still provide a good approximation of the "market potential" of a given technology, which is complementary to the "technological potential" which can be approximated with patent statistics.

4. Introducing the data and illustrating trends

Data sources

All the patent indicators presented in this paper are calculated using patent data from the European Patent Office (EPO)[xv] sourced from the EPO's Worldwide Patent Statistical Database – PATSTAT, edition Spring 2014.

All indicators are averaged at the level of 8 digit Y02-tag CPC sub-groups (e.g. Y02E10/1) and year, from 1991 to 2013.[xvi] The algorithm to assign Y-tags to patents largely builds upon IPC classes, which were introduced following the Strasbourg agreement in 1971. The data extraction does not go further back in time – although patent data are theoretically available from 1978 – because it is very likely that IPC classes themselves have taken some time to be fully operational (and used) after 1971, which would bias our patent counts towards more recent years. Y02 tags are assigned to climate change mitigation relevant technologies, which are also used to refer to as environmentally relevant technologies.

To compare our figures to a counterfactual, a 3% random sample across all technologies of approximately the same size (around 111 000 patent applications, compared to 135 000 applications in the Y02 sample) is constructed; in the following, the sample will be referred to as "the counterfactual".

4.1. Patent Counts

There are four subcategories of Y02 tags. Table 2 gives a quick overview of the classes and the size of the dataset under scrutiny. All numbers represent total single counts for the period 1991 – 2013; Annex A contains a detailed description of Cooperative Patent Classification (CPC) classes. Because the indicators are based on single counts (i.e. each y-tag assigned to a patent counts as one) and a patent can receive several y-tags, the figures below cannot be interpreted as proper patent counts, but rather as the prevalence of y-tags, and thus environmentally relevant technologies. The most important field is energy generation, transmission, and distribution (Y02E), with more than 64 000 patent applications over the sample period.

Also note that data for the years after 2011 are not reported in all following descriptive figures. Patent applications plummet after 2011, which makes quality indicators unreliable for circumscribed technology fields. The decline in patent numbers is an artefact of a significant lag in the data due to administrative processes, the fact that IP offices report to the EPO on a periodic basis and that the EPO database is updated once every 6 months.

In the counterfactual, patent application growth has taken off in the mid-90s and continued to grow at a relatively high pace. The recent downturn can be explained by the fact that there is a recording lag of a few years in the EPO database (even more so for granted patents).

However patent growth in climate change mitigation (CCM) technologies has been spectacular since the mid-1990s. Application numbers have basically quadrupled (again this is not a count of patents but a measure for how widespread environmentally relevant technologies are in applications). Using fractional counts, normalized to 100 in 1991, the picture looks very similar (see figure A.C.2, Annex C). This confirms that the single count method does not significantly bias counts when they are normalised, and it is a robust measure of prevalence of environmentally relevant technologies.

Interestingly, granted patents did not keep pace with the growth rate of the applications and only rose by around 50% since 2000. In fact the granted ratio (calculated as the number of granted patents over all applications) is almost identical between environmentally relevant technologies and the counterfactual and declining over time (see figure A.C.3. in the Annex C). This is important because it means that the fact that the observation that Y02 granted patents did not grow to the same extent as applications does not reflect a decreasing relative quality of patents related to the environment. There is a more systematic reason that prevails across all technologies and patents, which led to a decline in patent application success rates over time.[xvii]

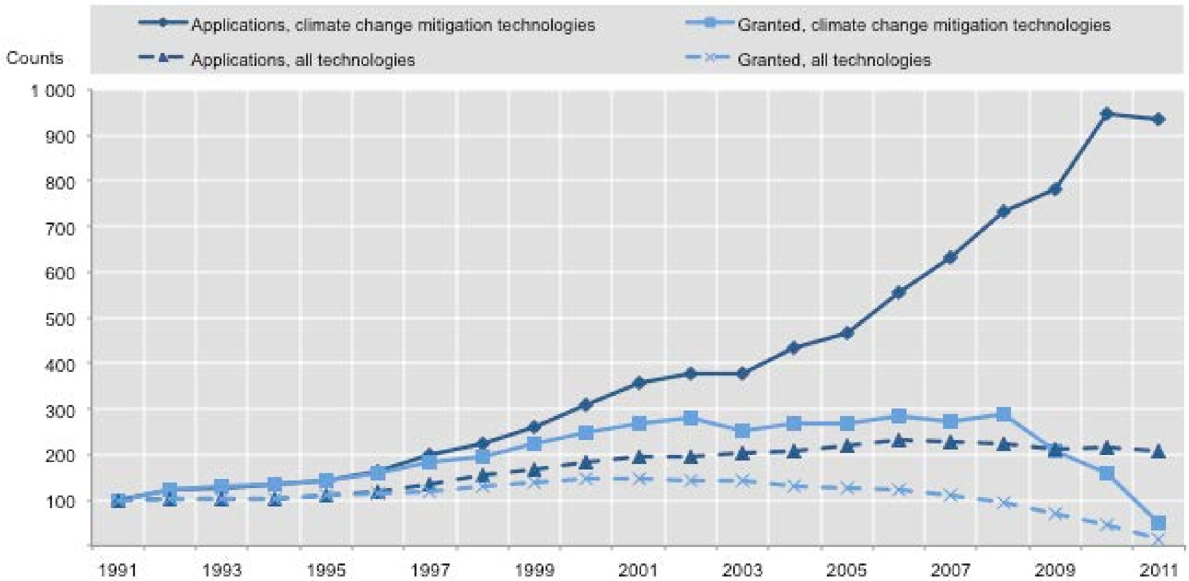
Table 2. Patent counts and Y02 tags

CPC class	Field	Number of 7 digit CPC classes	Number of patents	Granted
Y02	Technologies or applications for mitigation or adaptation against climate change	101	135 268	46 582
Y02B	*Related to buildings, e.g. including housing and appliances or end-user applications*	*39*	*23 444*	*7 154*
Y02C	*Capture, storage, sequestration or disposal of greenhouse gases*	*5*	*3 058*	*931*
Y02E	*Related to energy generation, transmission or distribution*	*30*	*64 281*	*18 743*
Y02T	*Related to transportation*	*27*	*44 485*	*19 754*

Source: OECD elaboration based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014.

Figure 3. Growth in patent applications and granted patents 1991-2011

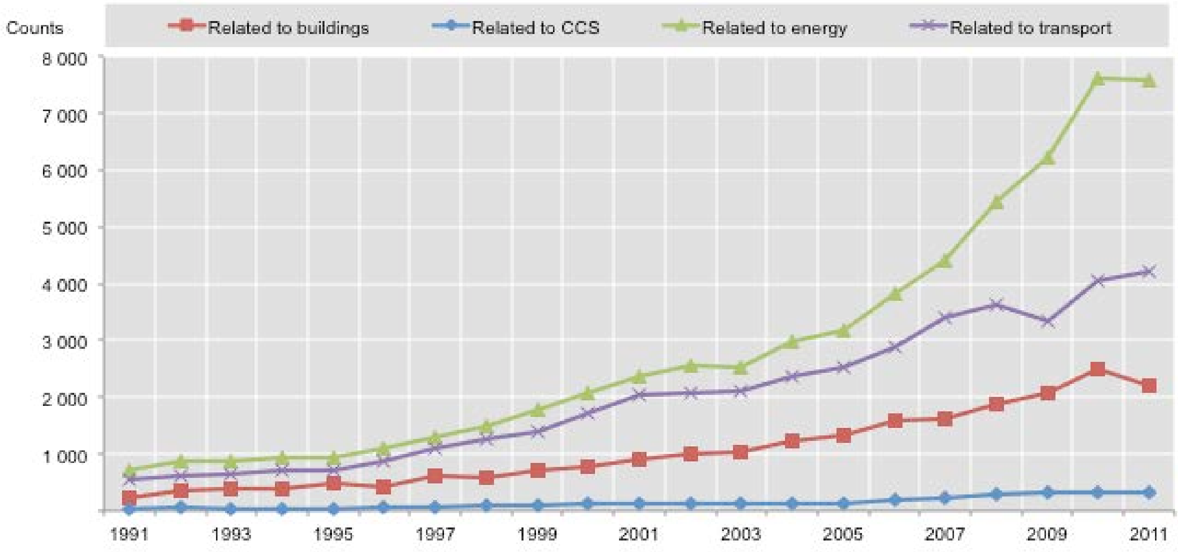
(normalised to 100 in 1991)



Note: the graphs report single counts per sample per year.
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014.

Further disaggregation of the picture into the four Y02 subcategories shows that growth in CCM related technologies mainly happened in the energy and transport sector. To some extent the technologies related to buildings contributed as well, but certainly CCS technologies only account for a very small share. However the CCS category is also by far the one with the fewest y-tags (see Annex C). This already gives a first indication that in the past 10 years the “blockbuster environmental technologies” can be found in the field of energy and transport. The picture for granted patents is very similar (see Annex C).

Figure 4. Growth in patent applications by 4-digit technology fields 1991 - 2011



Note: the graphs report single counts per each of the four sub-samples per year.

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014.

Patent Quality indicators

This section aims at identifying which technology fields have relatively high (and low) quality indicators, both relative to other Y02 fields and relative to the counterfactual. In addition, the aim of this paper being to identify ex-ante quality indicators that help predict whether a certain technology will be successful in the future, it is essential that these quality indicators are generally stable over time, i.e. that they can be interpreted as *technology specific characteristics*. This will then firstly allow ranking technologies with respect to their quality characteristics and, secondly, it will allow for the implementation of a pooled cross-section regression aimed at identifying which characteristics determine the flow of patent applications (and other dependents).[xviii]

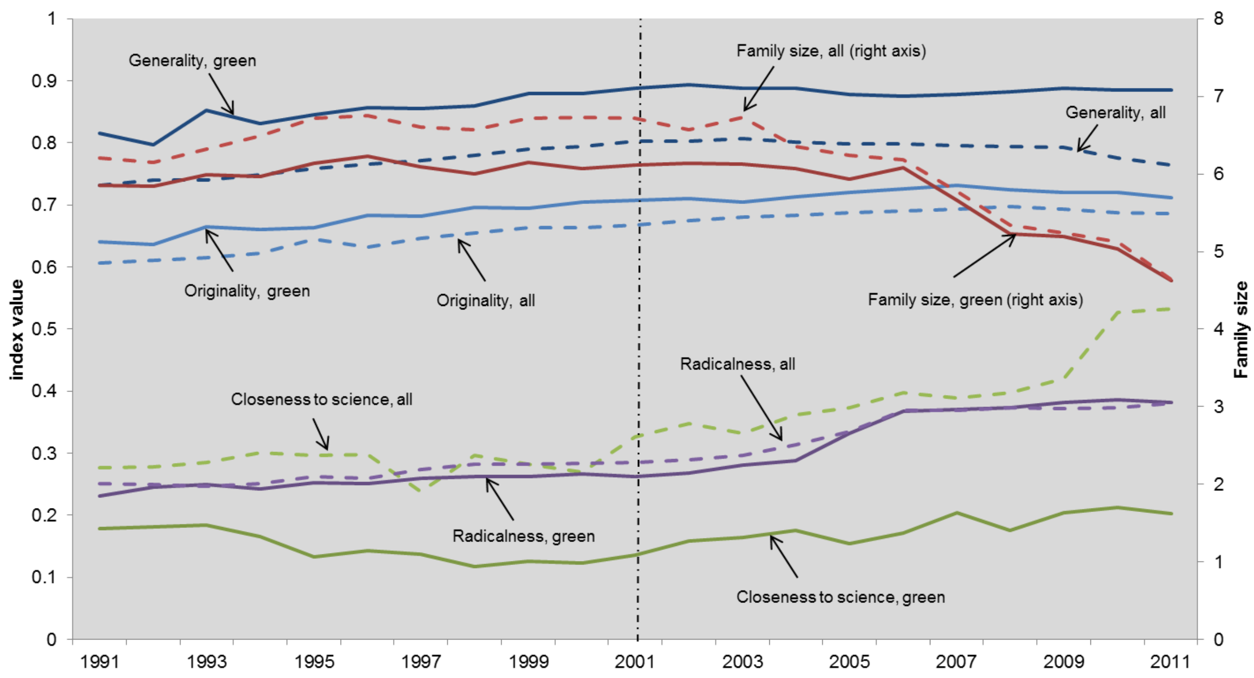
The visual inspection of time-variance uses the entire time period (1991-2011). This inspection is important because the analysis of the indicators is more meaningful if the quality indicators remain constant when a technology class takes off, i.e., a sharp increase in patent applications is observed. To illustrate the time-invariance, figures 6-10 plot all five quality indicators for (1) the entire Y02 sample in figure 5 and (2) for each of the 4 subcategories. This gives an idea of whether there are *trends over time* in the data. As noted, for purposes of comparison, the counterfactual is also included in these figures.

Looking at the full Y02 sample in figure 5 (solid line), there are two quality indicators for which there is a significant difference between the Y02 classes and the counterfactual: a) the industrial generality index, where the values for the Y02 are much wider (indicating wider diffusion across firms in different sectors; and, b) the closeness to science indicator where the counterfactual has considerably higher values.

In terms of temporal variation from 1991 to 2000 – the time period over which the indicators are constructed – there is no discernible time trend. Most of the quality indicators seem to show a slight upward trend but it is minor and in some cases even reversed later in the period. The indicator of family size shows a clear dip after 2006, however much of this might be an artefact due to the lags inherent in measures based upon duplicate patent applications.

The counterfactual sample shows a similar albeit not identical picture (dotted lines). Certainly from 1991 to 2000 the picture is almost identical for all indicators except industrial generality, where Y02 technologies are consistently higher. After 2001 the counterfactual sample shows a clear trend becoming closer to science, whereas this indicator is remarkably stable for Y02 technologies. This is interesting because one would expect closeness to science to be negatively related to commercial applicability (i.e. if a technology refers a lot to academic research, it is further away from the market). However the counterfactual sample does not become less applicable over time.

Figure 5. Synopsis: the five quality indicators 1990-2011, Y02 vs. counterfactual

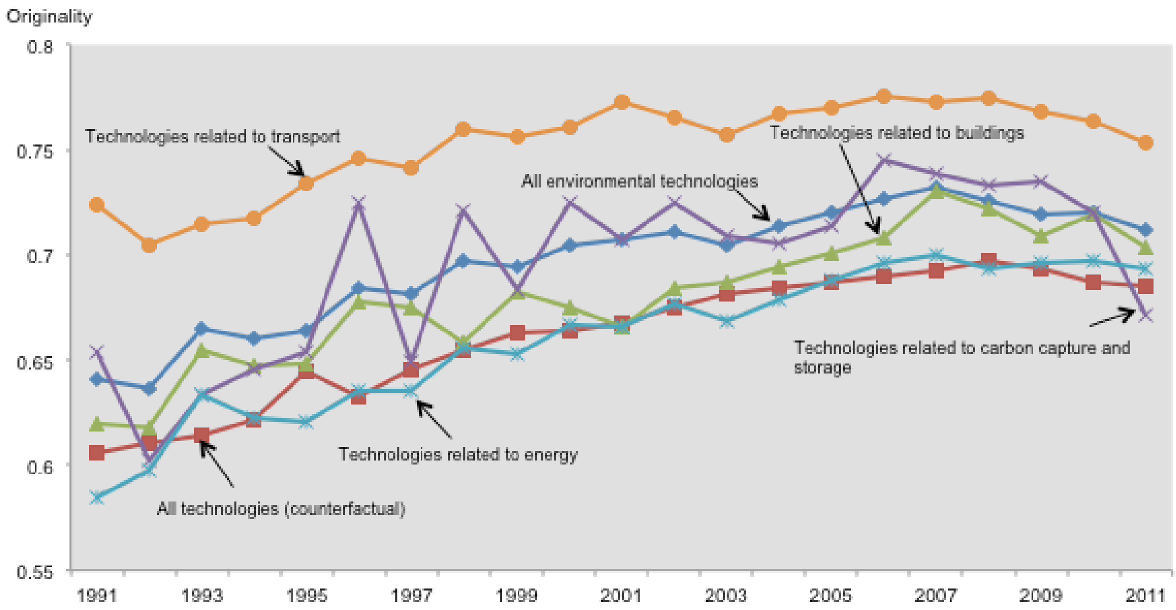


Note: Each of the five quality indicators is computed for the entire Y02 (climate change mitigation technologies) sample and for the counterfactual (all) sample for each year. The solid lines show the evolution of the quality indicators for the Y02 sample over time, the dotted lines show the evolution of the quality indicators for the counterfactual sample over time. By definition all indicators except for family size (right axis) are bounded by 0 and 1. The dotted black line indicates the end of the time period for which the quality indicators are computed for the cross-section regressions.

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

In order to get a better idea of differences within the Y02 classes, the sample is split into the four subsamples that were discussed previously (where energy and to a lesser extent transport are the most prevalent ones) in figures 6-10. For the time 1991-2000 the quality indicators are again very stable but there may be trends in more recent years. The technologies related to carbon capture and storage show greater volatility. This might be due to relatively low counts (See Figure C1). In what follows six time series (counterfactual, all Y02 technologies, and the 4 subcategories) for each quality indicator are shown, in order to briefly discuss patterns and time trends where discernible from a visual inspection.

Figure 6. Originality, 1991-2011

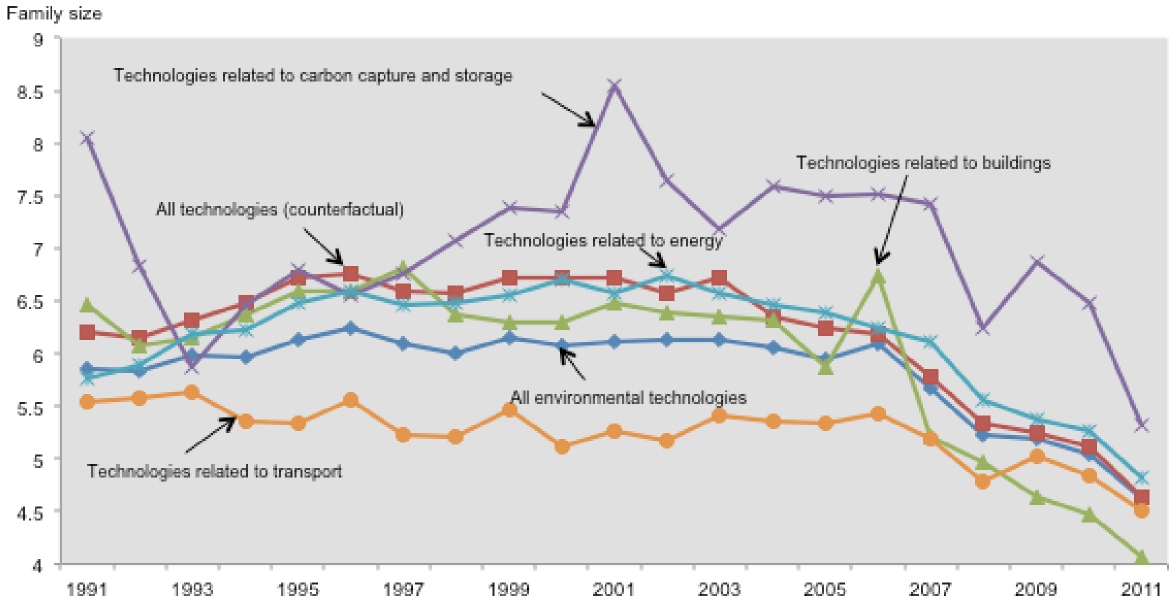


Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

Originality: Transport technologies are persistently more original than the average counterfactual and so is the measure of aggregate Y02 technologies. Energy technologies, on the other hand, seem not to be different from the average technology (both in terms of levels and trends). The graph also shows that there is a general upward trend over time affecting all technologies.

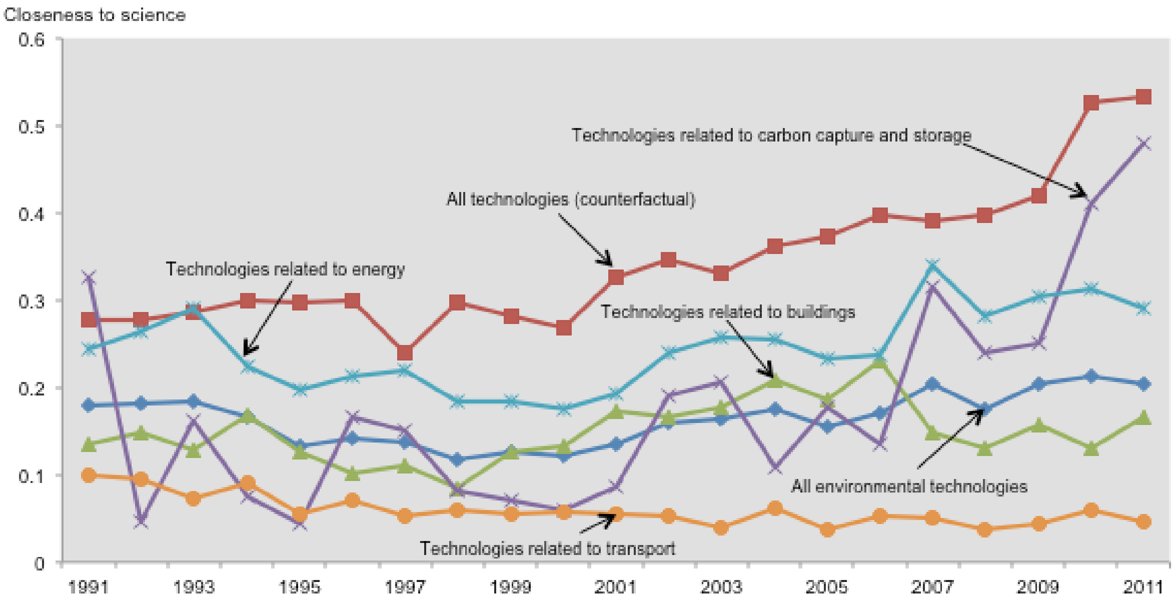
Family size: Again transport related technologies differ slightly from the other technologies in terms of family size, as patents in this field are filed at fewer patent offices that the average.[xix] CCS technologies, on the other hand, show a higher family size. This may depend on the fact that CCS technologies may have wider applicability across countries and higher economic value, whereas inventions in transport technologies can be more constrained by local regulation or have lower economic value. The same is found for building technologies, which has declined below the Y02 average as well. The trend is not clear but goes towards a geographically more narrow application base, i.e. smaller family size over time. Energy technologies are similar to the counterfactual, both in terms of level and trends.

Figure 7. Family size 1991-2011



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

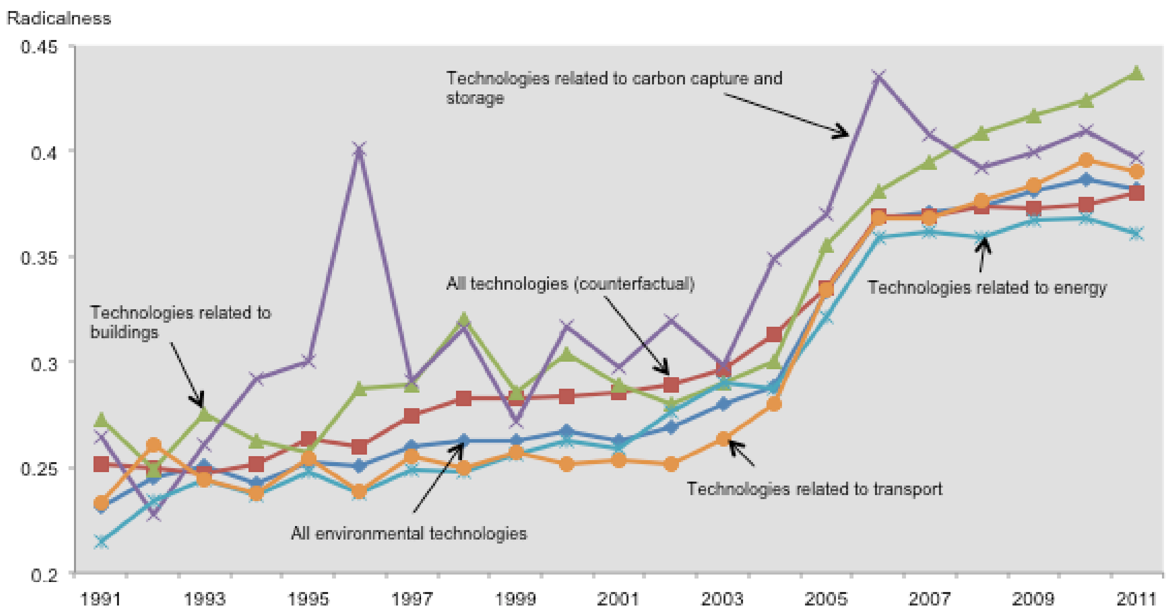
Figure 8. Closeness to science, 1991-2011



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

Closeness to science: The counterfactual values for this indicator are well above those for all Y02 technologies. There is also a trend for the average technology to involve more citations to science publications (i.e. reflecting increasing closeness to science). For the Y02 technologies, this trend is only mirrored in CCS technologies, however not for the average green technology. Transport technologies are well below all others with less than 10% of all patent citations going to literature in science.

Figure 9. Radicalness, 1991-2011



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

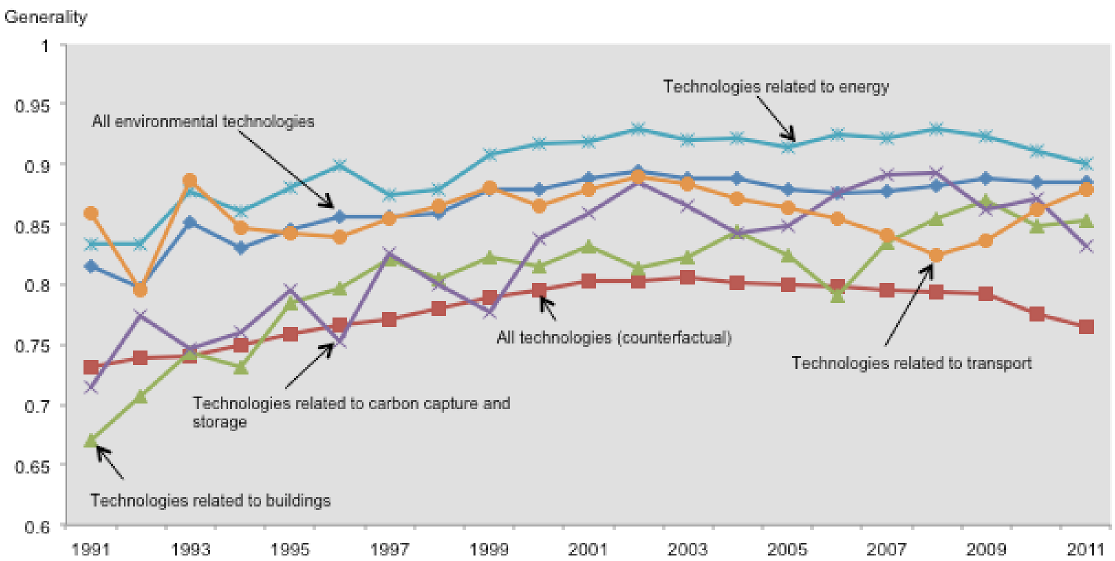
Radicalness: Except for a few outliers, the picture for this indicator is very homogenous across all subcategories and technological classes. There is an increase around 2004 that is correlated across all technology classes; technologies are becoming more radical.[xx]

Industrial generality: In this case the values for the counterfactual technology are below the Y02 classes, i.e. green technologies are more general than their counterfactual counterpart. There are quite remarkable differences in levels between subcategories. This is especially the case for technologies related to energy and also to a lesser extent for technologies related to transport (which are the two fields that experienced the highest growth in applications).

In order to test for the reliability of the use of our quality indicators in empirical analysis, Figure 11 presents the coefficient of variation (standard deviation divided by the mean) for all quality indicators, which allows for the identification of possible variation around the mean across time.

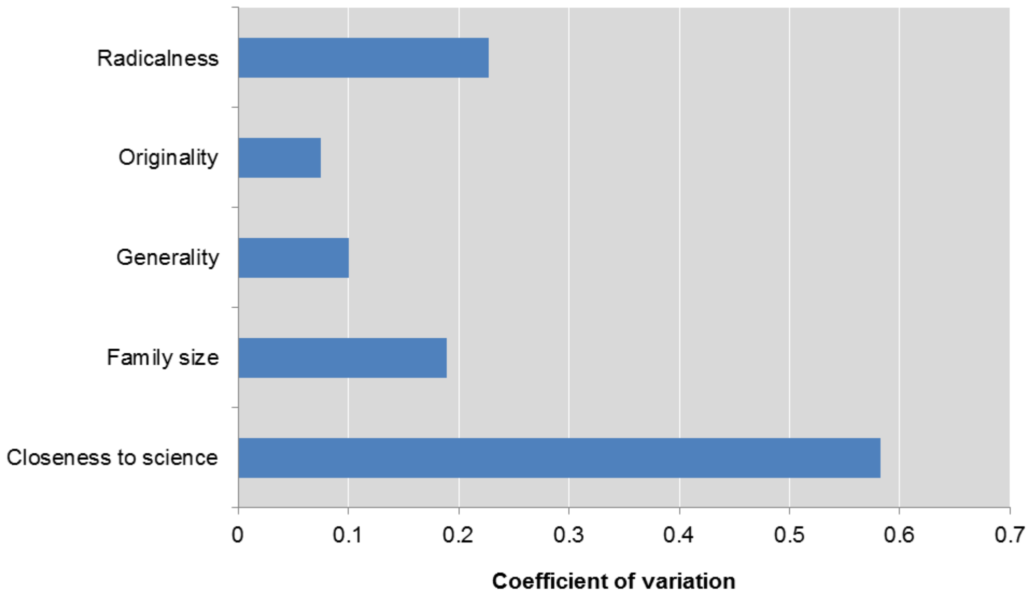
Calculating the coefficient of variation (CV) (standard variation over the mean) confirms that the indicators only show very little variation around the mean over time. For all indicators the CV is smaller than one and, except in the case of the indicator of closeness to science, it hovers around 0.1 – 0.2, thus indicating constant values across time. As closeness to science is a ratio, the variation is naturally bigger as it cumulates the variation of two variables (citations to non-patent literature and total backward citations).

Figure 10. Industrial generality, 1991-2011



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

Figure 11. Coefficient of variation for all quality indicators across the entire time period



Note: the coefficient of variation is calculated as the ratio of standard deviation over the mean.

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and on SDC.

Finally the analysis of the correlations between the quality indicators, which are all used as independent variables in the regressions of this paper, shows there is almost no correlation among them. This is true using the standard Pearson coefficient of correlation and also using the Spearman rank coefficient of correlation. Table 3 shows correlations between year and CPC class and table 4 shows correlation only between CPC classes (aggregated means over time) for different sample sizes.

For all specifications, the indicators are independent from one another, which first means that they reflect different patent characteristics; second, it also confirms that they can be safely used as explanatory variables in multivariate regression analysis without incurring risk of quasi-multicollinearity.

As the only exception, it must be noted that Originality and Radicalness show a correlation larger than 0.5 in all specifications. This is due to the fact that the two indicators indeed capture similar characteristics. An original patent relies on many fields of technology (breadth), which by default makes this patent more likely to be radical, i.e. dependent on many technologies from fields other than its own. They are therefore included alternatively in any given regression, as also reflected by hypotheses H1a and H1b for the indicators Originality and Radicalness.

Table 3. Correlations of quality indicators between year and CPC classes (obs = 1777)[xxi]

	Originality	Radicalness	Industrial Generality	Closeness to science	Family size
Originality	1.00				
Radicalness	0.51	1.00			
Industrial Generality	0.15	0.01	1.00		
Closeness to science	-0.17	0.03	0.01	1.00	
Family size	0.06	-0.04	-0.10	0.06	1.00

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table 4. Correlations of quality indicators between CPC classes (obs = 101)xxii and in brackets for CPC classes with >200 applications (obs = 63)

	Originality	Radicalness	Industrial Generality	Closeness to science	Family size
Originality	1.00 (1.00)				
Radicalness	0.58 (0.65)	1.00 (1.00)			
Industrial Generality	0.14 (0.23)	-0.07 (0.15)	1.00 (1.00)		
Closeness to science	-0.17 (-0.07)	0.12 (0.10)	0.03 (0.20)	1.00 (1.00)	
Family size	0.07 (-0.16)	0.10 (-0.11)	-0.22 (-0.20)	0.29 (0.42)	1.00 (1.00)

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Ranking technologies by quality indicator

In reality providing advice as to whether “energy” or “building related” technologies are more promising targets of support is not very useful for policy makers. The value of quality indicator measures is likely to be greater for lower levels of aggregation - policy makers would want information on which technology among energy related technologies (i.e. fuel cells vs. photovoltaics) and also which technology across different fields (i.e. photovoltaics vs. semiconductors) is the most promising for the future.

The analysis therefore distinguishes between different technologies *within* energy or *within* transport in the following sections to identify promising technologies. This is done by displaying the mean across time by indicator for all CPC classes with more than 1 000 single counts in figure 12 (henceforth, this selection is referred as the *>1000 subsample*). This allows for the construction of technology-level measure for each indicator separately and identifies the ranking of technologies at a detailed level. The 10th and 90th percentile for each indicator and each technology class are also calculated, which gives a nice picture of the *within technology class* variation of a quality indicator (figures A.C.4 – A.C.8). All figures also include the total (of the >1000 Y02 subsample) and the counterfactual.

The restriction to the >1000 sample is mainly due to the fact that the quality indicators are calculated by year and then averaged across time. If a CPC class does not have a sufficient number of patent applications per year the calculation of quality indicators can become much more volatile. The restriction does not entail a large sample reduction: the total sample with 101 CPC classes contains 135 268 patent counts, whereas our subsample with the 32 CPC classes containing more than 1 000 patent applications still covers 117 440 applications, which is 87% of all patent applications in our sample. Looking at table 5 it is also possible to see that restricting the sample does not lead to a sample bias. While it is true that the >1000 restriction increases the relative importance of energy and transportation patents in the total, the change does not appear to be so important as to bias the findings of the descriptive analysis.

Table 5. Subcategory representation in the subsample

	Full sample	>1000 subsample
Buildings	39%	28%
CCS	5%	3%
Energy	30%	34%
Transport	27%	34%
Total	100%	100%

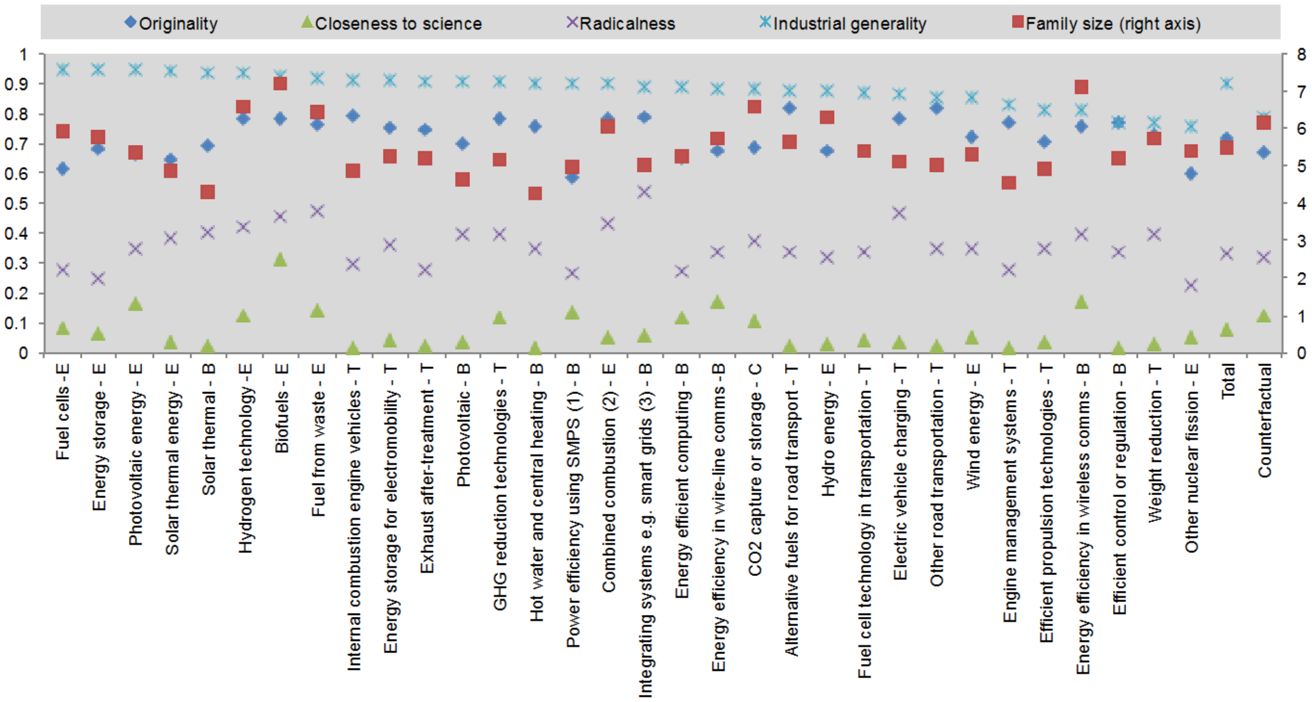
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

The ranking of all technologies (CPC classes) with more than 1 000 patent applications between 1991 and 2011 is also reported. In Annex D these technologies are ranked within the entire Y02 sample for the five quality indicators.[xxiii] The top 10 ranks are highlighted in red and it is possible to see that for all indicators except Industrial Generality the subsample of fields with more than 1 000 patents only covers up to four of the top 10 ranks. Interestingly however there is a full coverage of all top 10 technologies in the case of the indicator of industrial generality.

Furthermore, in figures A.C.4 – A.C.8 in Annex C, which plot the means and the 10th and 90th for each quality indicators for all 7 digit CPC classes in the >1000 subsample, it can be seen that, with few exceptions, variation across time on this more disaggregated level is still is quite low. It can first be concluded that the procedure of constructing the average of quality indicators for the period 1991-2000 per CPC class and taking them as technology characteristics for the cross-sectional regression is a promising procedure. Second, the descriptive figures A.C.4 – A.C.8 are very useful to identify prospective “champion” technologies once we have identified quality characteristics which have a statistically significant influence on subsequent developments.

The technologies in figure 12 are sorted by the indicator of Industrial Generality because the empirical analysis (presented in Section V) finds this indicator to be the one determining growth in patent applications. It is worth highlighting a few very promising technologies. Going through the top 10 technologies for Industrial Generality and anticipating some of the results in section 6, it is possible to see that **Photovoltaic Energy Generation (Y02E10/5)** ranks high in Industrial Generality (3rd) and Closeness to Science (13th) and low in Radicalness (61th), which are all characteristics that point to high patent applications in the future (also high levels of granted patents). **Hydrogen technology (Y02E60/3)** emerges as another interesting case. Ranking 6th in Industrial Generality, it also ranks in the top 25 in all other indicators. This is not necessarily pointing to high future application counts but it indicates a high potential of the technology to attract risk finance – characterising as a risky but potentially very successful technology. **Biofuels (Y02E59/1)** ranks in the top 20 for all indicators too. Ranking 7th in Industrial Generality, 3rd in Closeness to Science and 5th in Family Size it is in the top 10s for three indicators. Again this points to a risky but potentially very successful technology. Remarkably all of these technologies are located in the subcategory E (technologies related to energy generation, transmission or distribution). Interestingly, some of the most widespread technologies – and widely subsidised – such as hydro energy, wind energy or nuclear fission show lower values than the average technology.

Figure 12. Means 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



Note: (1) Switched-mode power supplies (2) e.g. heat utilisation (3) In power network, communications and IT.
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

5. Econometric analysis: model and results

As a complement to the descriptive evidence presented so far, this section describes the results of the econometric analysis, aimed at isolating some patent quality attributes which are associated with later technological and market diffusion. The analysis is preliminary in nature, but is intended to illustrate how such measures can be used in a manner which is informative for policy makers.

The estimated model is the following:

$$y_{k,2001-2010} = \alpha + \bar{\beta} Q_{k,1990-2000} + \bar{\gamma} X_{k,1990-2000} + \varepsilon_k \qquad (1)$$

where k indexes CPC technological classes at the 9-digit level of disaggregation (e.g., Y02B10/10; i.e., one level further with respect to descriptive statistics), Q is the set of five quality indicators presented above: Industrial Generality, Radicalness, Originality, Closeness to Science, Family Size; and y is one of the following measures of technological or market diffusion: total number of patents granted, total number of patent applications, commercial applicability index, and risk finance attractiveness.

Given the conceptual similarity of the Originality and Radicalness indicators, those are included separately and thus the regressions are run twice for each of the four dependent variables. X is a matrix of three control variables: a proxy for the technology maturity, the total number of patents in the first period, and the average number of backward citations. Maturity is measured for each technology by the average of the maximum citation lag of all the patents within that technology. The second variable controls for a “size-effect” which may affect the second period growth. In addition, it also filters out some of the potential noise which may affect those quality indicators which are based on concentration measures. For example, there could be a correlation between the industrial generality index and the amount of patents from which the indicator is generated. The inclusion of this control ensures that the coefficients on the generality index are not biased. In a similar vein, the third control variable is included to estimate more precisely the effect of the quality indicators, by controlling for a more general indicator of quality based on backward citations.

The model is estimated by least squares (LS), with robust standard errors; all the variables are expressed in logarithmic form. Given that the relative weight in terms of patents of CPC technological classes is extremely heterogeneous – with 62 classes out of 266 included in the regression analysis having less than 10 patent applications before the 2000, and 38 classes having more than 200 – the regressions are also analytically weighted by the total number of patent applications in the 1990-2000 period.

Importantly, the dependent variables are calculated on patents applied for in the year 2000’s, while the quality indicators are calculated based on patents applied for from 1990 to 2000. The choice is driven by the visual inspection of the series of patent applications, which shows that the turning point of the most successful technological classes is around year 2000. Robustness tests setting different turning points (from 1998 to 2002) produce similar results.

The first set of results is shown in Table 6. As it is possible to see, the indicator of Industrial Generality is the only one that is always robustly associated with later technological or market diffusion, i.e., technologies that are used by a differentiated set of industries are also those which: experience higher increase in granted patent and patent applications; in which applicants tend to be more often private companies; and which attract more risk finance.

The coefficient on Radicalness and Closeness to Science are also significant in most regressions. More radical technologies in the 90’s show a relatively slower technological diffusion over the year 2000’s and also tends to be less popular among private patent applicants. Closeness to Science is positively associated with the number of later patent applications, but not with the number of granted patents; technologies which are closer to science are also less frequently patented by private business applicants. Importantly, both Radicalness and Closeness to Science are positively associated with later attractiveness of risk finance. This latter result may look apparently surprising but it is actually consistent with findings from Criscuolo and Menon (forthcoming), who show that the two indicators are positively associated with the *probability* of getting funded (although radicalness is negatively associated with the *amount* of funding received from venture capitalists and other early stage investors). Finally, the negative coefficient on the closeness to science indicator in the regression with *applicability* as dependent variable is not surprising, as patents which are closer to science are more likely to be developed by public research institutions.

Finally, the results from the regression with the Originality indicator (Table 7) suggest that more original technologies are more frequently patented by private business applicants; the indicator turns out not to be significant in all other regressions.

Among the control variables, the number of patent applications in the first period has the expected positive coefficient. The maturity proxy is, however, not always significant, with the exception of the regression of the commercial applicability indicator, in which it has a negative coefficient. While this result is tangential to the main focus of the paper, a speculative explanation of the results may rest on the stronger inertia of non-private research institutions, which leads them to focus on more mature technologies. Conversely, the business sector may show faster reactivity to new technological fields.

Table 6. The predictors of technological diffusion in 2001-2010

Dependent variable:	Patents granted 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Applicability 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	4.860***	4.731***	0.0729**	18.82***
	(0.338)	(0.321)	(0.0332)	(2.837)
Radicalness	-3.285***	-1.766**	-0.124	25.60***
	(0.862)	(0.824)	(0.0833)	(9.689)
Closeness to science	0.465	2.466**	-0.436***	32.67***
	(1.034)	(1.079)	(0.149)	(9.927)
Family size	-0.00976	0.0430	-0.0144**	1.738***
	(0.108)	(0.0947)	(0.00659)	(0.512)
Maturity	-0.00925	0.0852	-0.134***	0.973
	(0.316)	(0.286)	(0.0227)	(2.366)
Patent apps. 1990-2000	0.405***	0.420***	0.00457	2.453***
	(0.0828)	(0.0706)	(0.00334)	(0.391)
No. of backward cits.	0.225*	0.207*	-0.00916	0.937
	(0.123)	(0.117)	(0.0114)	(1.090)
Constant	1.371	1.300	1.698***	-32.10***
	(1.560)	(1.395)	(0.0981)	(12.20)
Observations	266	266	266	266
R-squared	0.795	0.798	0.421	0.449

*All regressions are estimated with Least Squares (LS) and are weighted by the no. of patents in the period 1990-2000 in the CPC category. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table 7. The predictors of technological diffusion in 2001-2010 (with Originality)

Dependent variable:	Patent granted 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Applicability 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	4.928***	4.787***	0.0608*	18.42***
	(0.398)	(0.350)	(0.0316)	(2.877)
Originality	0.705	0.114	0.237***	-7.406
	(1.033)	(0.888)	(0.0748)	(8.577)
Closeness to science	0.888	2.608**	-0.353**	28.75***
	(0.968)	(1.055)	(0.145)	(10.16)
Family size	-0.0637	0.0148	-0.0172***	2.164***
	(0.105)	(0.0930)	(0.00533)	(0.539)
Maturity	-0.00955	0.0686	-0.121***	0.857
	(0.297)	(0.268)	(0.0227)	(2.312)
Patent apps. 1990-2000	0.473***	0.454***	0.00948***	1.903***
	(0.0850)	(0.0704)	(0.00322)	(0.310)
No. of backward cits.	0.176	0.193	-0.0211*	1.411
	(0.123)	(0.118)	(0.0116)	(1.046)
Constant	0.472	1.010	1.509***	-23.69*
	(1.436)	(1.288)	(0.115)	(13.29)
Observations	266	266	266	266
R-squared	0.781	0.793	0.453	0.428

*All regressions are estimated with Least Squares (LS) and are weighted by the no. of patents in the period 1990-2000 in the CPC category. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

5.1. Robustness

The consistency of the discussed results has been assessed by a number of robustness checks, the results of which results are reported in Appendix F.

CPC group dummies: the regressor list is augmented by a dummy variable for each of the four main CPC group listed in the Y02 category (identified by the first letter following the Y02 coding, i.e., B, C, E, and T). Results are unaffected.

One indicator at the time: although patent quality indicators are generally not too strongly correlated among them, to be on the safe side the results are tested by including them in the regression one-by-one, in order to check whether quasi-multicollinearity may jeopardize the results. The test confirms that the latter is not an issue, as one-by-one coefficients generally convey the same messages of the multivariate regressions, with the notable exception of the coefficients on closeness to science, which loses its significance. Results on the latter variables should therefore be interpreted as conditional on the other independent variables.

Controlling for public R&D spending: national policies may affect the development path of different technologies. Our dataset is not aggregated by countries, however to the extent that technologies are unevenly distributed across countries, the omission of information on national policies may lead to inconsistent estimates. Although data on country- and technology-specific policies are extremely hard to find, it is possible to build a proxy based on public R&D spending by country, year, and industry (IEA, 2013). As the classifications used for the R&D data corresponds closely with the CPC technological breakdown used in this study, the two are matched manually based on label similarity (see Annex B). However, the R&D data classification is more general than the 8-digit CPC breakdown, thus the R&D data are grouped into 17 categories; to each CPC technology is attributed the total value of the R&D category it is matched to. Probably also as a consequence of that, the results are inconclusive on the role played by public R&D expenditures. However the coefficients on quality indicators are again unaffected.

Estimation on sub-sample of technological classes with at least 20 patent applications in the period 1990-2000: although the weighting is already limiting the influence of potential outlier with very few patents in the first period, the consistency of the results is checked by replacing the weighting in this alternative way. Once again the main results are extremely similar to those presented in the baseline estimates.

Controlling for the applicant firms' structure: although the sign and the magnitude of the industrial generality variable is remarkably robust over a number of different specifications with alternative dependent variables, to be on the safe side an additional robustness test is implemented. Specifically, one may be concerned that the coefficient on the indicator absorbs the effect of other, unobserved characteristics of the applicant firms recorded in the HAN database. In order to address this concern an additional control variable was constructed, measuring the share of young firms in the total number of firms by 9-digit CPC technology. Although the new variable is often significant, coefficients on the industrial generality variable are almost identical to those reported in the main specifications.[xxiv]

Controlling for the weight of multinational enterprises (MNEs): technologies in which MNEs play a preponderant role as sources of invention may be thought to both more sectorally diversified, as well as having more potential for future growth. While, the inclusion of a variable measuring the number of patents in the first period should control for this effect, in order to ensure that the results are robust the regressions are also estimated including an additional control variable measuring the share of patents in each technology that are applied for by multinational affiliates.[xxv] Again, the main results are unaffected, although the additional control is significant and positive in some specifications.[xxvi]

6. Conclusions

This paper should be seen as a preliminary exploration of how to meet an important policy challenge – a method for the identification of future breakthrough technologies. First, the report tries to predict the future technological and market success of technologies using existing data, against a backdrop of fundamental uncertainty. Second, the analysis combines data sources that have not yet been combined in analytical work (e.g. CleanTech data on venture capital, Orbis data on firm sectors, and OECD/EPO data on patents), using indicators that have only been developed very recently and have not been applied in other publications. Third, the Y-tags used to define technologies are a very new development, and represent a considerable effort on the part of patent examiners at the EPO.

In essence this paper tries to provide guidance on how to target incentives in a technologically-discretionary manner. This is a hazardous exercise, and there is little doubt that the backbone of innovation policy should be technology-neutral (i.e. protection of intellectual property, investment in basic research, fostering collaboration between different actors in the innovation process, etc.). However, as a complement to this general and technology-neutral policy framework, in practice many governments are providing discretionary incentives for specific technologies, and this report seeks to explore in a very preliminary manner how such choices can be made.

The intention is to try and uncover attributes of inventions (as reflected in patent data) which serve as "leading indicators" of subsequent technological and market development. The objective is to identify a methodology which can provide guidance to policymakers in the choices made with respect to public support for different technological fields. The case of “climate mitigation” technologies serves as the field examined. However, the analysis is relevant for other technology fields as well, and this paper should be seen as a case study application.

Of the different quality attributes, the role of industrial generality emerges as being particularly important. Originality and radicalness have more ambiguous results, with signs changing depending upon the dependent variable applied. However, such ambiguity is consistent with previous findings. The measure of closeness to science has positive and significant effects, except in the case of our measure of commercial applicability. This is not surprising. In summary, the empirical analysis suggests that the only robust indicator of success for technologies is industrial generality. In every specification a more “general” technology predicts a growth in patent applications, granted patents, private industry propensity to patent and risk finance attraction.

Further work is clearly necessary to refine this analysis. First, it is important to control for factors for which data is not readily available – most notably national policies. Second, it is important to control for government policy interventions that favour certain technologies. The use of Research, Development and Design (RD&D) data is a preliminary step in this direction, but it could be applied for a larger set of countries and a wider range of technologies.[xxvii] Third, there are important methodological challenges including problems of endogeneity. Fourth, the work could be applied in other policy domains where policy support is significant (i.e. medicines) or fields where technological change is rapid and potentially economically significant (e.g. nanotechnology, for which it was also discussed to create a specific “nano” Y-tag transversal CPC group: see Scheu et al., 2006). Fifth, the role played by the larger private companies in spurring sudden developments of specific technologies could also further analysed. Finally, taking into account the temporal dimension in a more detailed way – notably by the means of a panel regression – may disclose interesting information on the “turning points” of fast-growing technologies. This would in turn allow for a characterization of key actors – countries, companies, inventors – in removing bottle-necks and in unleashing the full potential of a given technology.

It has been emphasised that providing discretionary support is a hazardous exercise, and so even if further work allows for the identification of indicators which appear to be reliable leading indicators to future technology and market developments, there will still be great uncertainty. As such, as a concluding remark it is important to bear in mind the suggestions of Rodrik (2004) who emphasises the importance of ‘trial-and-error’ and the establishment of transparent and unambiguous exit mechanisms. In fact, the government needs to exit not only if a technology proves to be unsuccessful, but also if a technology proves to be successful enough to be driven by private actors. Determining this threshold of private involvement is another research question.

NOTES

[i] Graduate Institute, Geneva.

[ii] Directorate for Science, Technology and Innovation, OECD.

[iii] For the purposes of this paper the construction of the quality indicators differs somewhat from those presented in Scquicciarini et al. (2013). In addition, some new measures are proposed. However, the measures adopted draw heavily upon this paper.

[iv] The authors wish to thank Chiara Criscuolo, Luca Marcolin and Mariagrazia Squicciarini for useful comments, Hélène Dernis for help with patent data, and Isabelle Desnoyers-James for help with data visualisation.

[v] And thus a high proportion of option values in total economic values on both sides of the equation.

[vi] In the early 1990s the Energy Model Forum's project involving comparison of mitigation costs from 12 different models, three backstop technologies were shale-based liquid synthetic fuels, carbon-free liquid fuels and "carbon-free" electricity. See Hoeller et al. (1992) for a discussion.

[vii] For a discussion of the relative merits of different measures see OECD (2008) and Johnstone, Haščič and Watson (2011).

[viii] The economics of patents literature has however uncovered interesting correlations of quality attributes with patent market and scientific value, with number of forward citations and family size being the most well-known example. Future work may use these indicators as proxy for market or technological development of patent classes.

[ix] For the purposes of this paper the construction of the quality indicators differs somewhat from those presented in SDC. In addition, some new measures are proposed. However, the measures adopted draw heavily upon this paper.

[x] However, this may have negative consequences for attractiveness to risk finance, one of the measures used to assess the subsequent market success of different technologies.

[xi] The patent generality index is traditionally calculated as the Hirschman-Herfindahl Index (HHI) of the distribution of citations received (forward citations) across technology classes (IPC) (cf. SDC).

[xii] The generality index is applied for a longer period of time than the coverage of the HAN database. As such it is assumed that the industry classification of firms did not change over the entire period of the estimation sample. Similarly, assignees of patents applied for at the beginning of the period may have ceased operating before 1998, therefore the matching ratio for older patents is also lower. Although this should not systematically bias the index in a precise direction, its value is surely noisier at the beginning of the period examined.

[xiii] "The objective of the sector allocation methodology is to allocate each patentee to one of the following sectors: (a) individual (private) patentee (b) private business enterprise (c) government (agency)

(d) university/higher education (e) private non-profit. This classification shows similarities with the existing sector classification developed by OECD in the context of conducting surveys on research and development, as outlined in the Frascati Manual (2002)." (p. 16). See http://epp.eurostat.ec.europa.eu/cache/ity_offpub/ks-ra-11-008/en/ks-ra-11-008-en.pdf.

xiv We thank Chiara Criscuolo for providing us with the aggregate indicator. The dataset is described at www.cleantech.com.

xv PATSTAT database, simple queries available online at www.espacenet.com.

xvi Depending on the question we vary the degree of aggregation of the CPC classes (i.e. from 4 up to 9 digits).

xvii A lower grant ratio might signal the maturity of a given technology, as the novelty requirement is met less often.

xviii In a sense this is a sort of "development economics regression", where time-invariant characteristics are the independent variables.

xix This is plausible given the structure of the transport industry and the localisation of production and R&D facilities in relatively few countries, hence the need to protect inventions at a relatively small set of offices.

xx The generalised sharp increase around year 2004 in patent radicalness is an interesting finding. Examination of the reasons why this is so is left for future research.

xxi Spearman rank correlation coefficients (assume linear relation between variables and are robust with regards to outliers) show very similar results except for Closeness to science – Industrial generality where the correlation is 0.25.

xxii Again the results are very similar using Spearman rank correlation coefficients, except for a slightly higher correlation between Closeness to science and Industrial generality.

xxiii CPC classes are ranked within the total of all 101 CPC classes, not the subsample of those 32 CPC classes with >1000 patent applications.

xxiv Results available upon request.

xxv The variable measuring the MNE weight in a given technology is calculated on information on business group structures taken from Andrews, Criscuolo, and Menon (2014) which is in turn based on elaborations on the OECD ORBIS database. The reader should note that, due to data availability, the group definition is based on data for the year 2009. For the period 1990-2000, firms are classified as MNE affiliates if in 2009 they belong to a business group in which there are affiliates from at least two different countries. The level of approximation is therefore not negligible; however this is the only viable option, to the best of authors' knowledge.

xxvi Results available upon request.

xxvii More recent data (i.e. from around 2005) is more disaggregated, allowing for richer analysis in the future.

REFERENCES

Andrews, D., C. Criscuolo, and C. Menon. "Do resources flow to patenting firms? Cross-country evidence from firm-level data." *Economics Department Working Papers no. 1127, OECD.*

Bosetti, V., C. Carraro, R. Duval, and M. Tavoni, (2001). "What should we expect from innovation? A model-based assessment of the environmental and mitigation cost implications of climate-related R&D" in Energy Economics, Volume 33, Issue 6, November 2011, Pages 1313-1320

Bresnahan T. F and M. Trajtenberg (1995), "General Purpose Technologies. 'Engines of Growth'?", *Journal of Econometrics*, Vol. 65: 83-108.

Criscuolo, C., and C. Menon, (2014). "Environmental policies and risk finance in the green sector: cross-country evidence." *OECD Science, Technology and Industry Working Paper n. 2014/1, OECD: Paris*.

Criscuolo C., and C. Menon, (Forthcoming). "Do patents matter for venture capital investments? Evidence from the green sector." OECD, Working paper.

Daim, Tugrul et al. (2006) "Forecasting emerging technologies: Use of bibliometrics and patent analysis" in *Technological Forecasting and Social Change*, Vol. 73, pp. 981-2012.

Gal, P. N. (2013). Measuring total factor productivity at the firm level using OECD-ORBIS. *Economics Department Woking Paper* No. 1049, OECD Publishing.

Gonnard, E., & A. Ragoussis (2012). "The OECD-ORBIS database treatment and benchmarking procedures". *OECD Statistics Working Paper*, Forthcoming.

Griliches, Z. (1998). "Patent statistics as economic indicators: a survey". In *R&D and productivity: the econometric evidence* (pp. 287-343). University of Chicago Press.

Guellec, D. and B. van Pottelsberghe de la Potterie, Applications, grants and the value of patent, Economics Letters, Volume 69, Issue 1, October 2000, Pages 109-114,

Haščič, I., Silva J., and N. Johnstone (2015) "Patent Statistics for International Comparisons of and Analysis of Narrow Technological Fields" *OECD Science, Technology and Industry Working Paper*, OECD Publishing (*forthcoming*).

Haupt, R., M. Kloyer and M. Lange (2007) "Patent Indicators for the Technology Life Cycle Development" in *Research Policy*, Vol. 36, pp. 387-398.

Hoeller, P. et al. (1992) "New Issues, New Results: The OECD's Second Survey of the Macroeconomic Costs of Reducing CO2 Emissions" OECD Economics Department Working Paper No. 123.

IEA (2010). Technology Roadmap: Concentrating Solar Power. IEA, Paris.

IEA (2013). RD&D Budget. *IEA Energy Technology RD&D Statistics* (database).

Johnstone, N. and I Haščič (2013) “Policy Incentives for Energy and Environmental Technological Innovation: Lessons from the Empirical Evidence” in Encyclopedia of Energy, Natural Resource and Environmental Economics (http://dx.doi.org/10.1016/B978-0-12-375067-9.00061-9)

Johnstone, N. and I. Haščič and F. Watson (2011) ‘Methodological Issues in the Development of Indicators of Innovation and Transfer in Environmental Technologies” Annex A in OECD *The Invention and Transfer of Environmental Technologies*. OECD, Paris.

Lanjouw, J. O., Pakes, A. and Putnam, J. (1998), How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data. The Journal of Industrial Economics, 46: 405-432. doi: 10.1111/1467-6451.00081

Löschel, A. and M. Schymura (2013) "Modeling technological change in economic models of climate change: A survey" ZEW Discussion Paper 13-007, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

OECD (2001). *Using patent counts for cross-country comparisons of technology output*. Paris: OECD.

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

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

OECD (2012) *Energy and Climate Policy: Bending the Technological Trajectory.* Paris: OECD.

Rodrik, D. (2004), “Industrial Policy for the Twenty-First Century”, paper prepared for UNIDO.

Scheu M., V. Veefkind, Y. Verbandt, E. Molina Galan, R. Absalom, W. Förster (2006), “Mapping nanotechnology patents: The EPO approach”. *World Patent Information*, 28:3, September, pp. 204-211.

Shane, S. (2001), “Technological Opportunities and New Firm Creation”, *Management Science*, 47 (2), pp. 205–220.

Squicciarini, M., H. Dernis and C. Criscuolo (2013), "Measuring Patent Quality: Indicators of Technological and Economic Value", *OECD Science, Technology and Industry Working Papers*, No. 2013/03, OECD Publishing. (doi: 10.1787/5k4522wkw1r8-en)

Trajtenberg, M., R. Henderson & A. Jaffe (1997) “University Versus Corporate Patents: A Window On The Basicness Of Invention” in Economics of Innovation and New Technology, Volume 5, Issue 1, 1997

Veefkind, V., J. Hurtado-Albir, S. Angelucci, K. Karachalios, and N. Thumm (2012), “A new EPO classification scheme for climate change mitigation technologies”, *World Patent Information* 34 (2): 106-111.

ANNEX A. CLASSIFICATION OF TECHNOLOGIES (CPC CLASSES)

Cooperative Patent Classification (CPC)							
Y02 - Technologies or applications for mitigation or adaption against climate change							
Y02E - Reduction of Greenhouse Gases (GHG) emission, related to energy generation, transmission or distribution							
6 digit	Technology	7 digit	Technology2	8 digit	Technology3	9 digit	Technology4
Y02E10	*Energy generation through renewable energy sources*						
		Y02E10/1	*Geothermal energy*				
				Y02E10/12	*Earth coil heat exchangers*		
						Y02E10/125	*Compact tube assemblies, e.g. geothermal probes*
				Y02E10/14	*Systems injecting medium directly into ground, e.g. hot dry rock system, underground water*		
				Y02E10/16	*Systems injecting medium into a closed well*		
				Y02E10/18	*Systems exchanging heat with fluids in pipes, e.g. fresh water or waste water*		
		Y02E10/2	*Hydro energy*				
				Y02E10/22	*Conventional, e.g. with dams, turbines and waterwheels*		
						Y02E10/223	*Turbines or waterwheels, e.g. details of the rotor*
						Y02E10/226	*Other parts or details*
				Y02E10/28	*Tidal stream or damless hydropower, e.g. sea flood and ebb, river, stream*		
		Y02E10/3	*Energy from sea*				
				Y02E10/32	*Oscillating water column [OWC]*		
				Y02E10/34	*Ocean thermal energy conversion [OTEC]*		
				Y02E10/36	*Salinity gradient*		
				Y02E10/38	*Wave energy or tidal swell, e.g. Pelamis-type*		
		Y02E10/4	*Solar thermal energy*				
				Y02E10/41	*Tower concentrators*		

		Y02E10/42	*Dish collectors*		
		Y02E10/43	*Fresnel lenses*		
		Y02E10/44	*Heat exchange systems*		
		Y02E10/45	*Trough concentrators*		
		Y02E10/46	*Conversion of thermal power into mechanical power, e.g. Rankine, Stirling solar thermal engines*		
				Y02E10/465	*Thermal updraft*
		Y02E10/47	*Mountings or tracking*		
Y02E10/5	*Photovoltaic [PV] energy*				
		Y02E10/52	*PV systems with concentrators*		
		Y02E10/54	*Material technologies*		
				Y02E10/541	*CuInSe2 material PV cells*
				Y02E10/542	*Dye sensitized solar cells*
				Y02E10/543	*Solar cells from Group II-VI materials*
				Y02E10/544	*Solar cells from Group III-V materials*
				Y02E10/545	*Microcrystalline silicon PV cells*
				Y02E10/546	*Polycrystalline silicon PV cells*
				Y02E10/547	*Monocrystalline silicon PV cells*
				Y02E10/548	*Amorphous silicon PV cells*
				Y02E10/549	*organic PV cells*
		Y02E10/56	*Power conversion electric or electronic aspects*		
				Y02E10/563	*for grid-connected applications*
				Y02E10/566	*concerning power management inside the plant , e.g. battery charging/discharging, economical operation, hybridisation with other energy sources*
				Y02E10/58	*Maximum power point tracking [MPPT] systems*
	Thermal-PV hybrids				
	Wind energy				
		Y02E10/72	*Wind turbines with rotation axis in wind direction*		

		Y02E10/721	Blades or rotors
		Y02E10/722	Components or gearbox
		Y02E10/723	Control of turbines
		Y02E10/725	Generator or configuration
		Y02E10/726	Nacelles
		Y02E10/727	Offshore towers
		Y02E10/728	Onshore towers
Y02E10/74	Wind turbines with rotation axis perpendicular to the wind direction		
Y02E10/76	Power conversion electric or electronic aspects		
		Y02E10/763	for grid-connected applications
		Y02E10/766	concerning power management inside the plant, e.g. battery charging/discharging, economical operation, hybridisation with other energy sources

ANNEX B. MATCHING OF IEA RD&D DATA TO CPC CLASSES

We use IEA data on RD&D budget in 2012 USD prices, PPP adjusted, across all 27 – mostly European – countries (IEA, 2013). Unfortunately the country selection of the IEA RD&D data does not match the respective applications by country to the EPO or the country shares in our dataset of patent applications. But as both samples include mostly European countries, the data can be used together.

In a next step we match the RD&D data, which is classified along end-uses to the CPC classes. We lose only 2 CPC classes and 829 patent applications doing this and are able to match an average of 66% of RD&D budget across all years to CPC classes. We end up with 17 RD&D categories containing each several of the remaining 99 CPC classes. The reason why the RD&D data is fairly aggregated to only 17 categories is that the IEA data only starts disaggregated recordings around 2005, which means the necessary length of time series is only available on a more aggregated level. The following table shows the matching of CPC classes to RD&D categories.

Table B.1. CPC to RD&D matching (RD&D in columns, CPC classes in rows)

Res. and comm. buildings, appliances and equipment"	Transport	CO2 capture and storage	Photovoltaics	Solar thermal power and high-temp. applications	Wind energy	Ocean energy	Biofuels (incl. liquids, solids and biogases)
Y02B20/1	Y02E20/1	Y02C10/0	Y02B10/1	Y02B10/2	Y02B10/3	Y02E10/3	Y02E50/1
Y02B20/2	Y02E20/3	Y02C10/1	Y02E10/5	Y02E10/4	Y02E10/7		Y02E50/3
Y02B20/3	Y02T10/1	Y02C20/1					
Y02B20/4	Y02T10/2	Y02C20/2					
Y02B20/7	Y02T10/3	Y02C20/3					
Y02B30/1	Y02T10/4						
Y02B30/2	Y02T10/5						
Y02B30/5	Y02T10/6						
Y02B30/6	Y02T10/7						
Y02B30/7	Y02T10/8						
Y02B30/9	Y02T10/9						
Y02B40/1	Y02T30/1						
Y02B40/3	Y02T30/3						
Y02B40/4	Y02T30/4						
Y02B40/5	Y02T50/1						
Y02B40/7	Y02T50/3						
Y02B40/8	Y02T50/4						
Y02B40/9	Y02T50/5						
Y02B50/1	Y02T50/6						
Y02B50/2	Y02T50/8						
Y02B80/1	Y02T50/9						
Y02B80/2	Y02T70/1						
Y02B80/3	Y02T70/3						

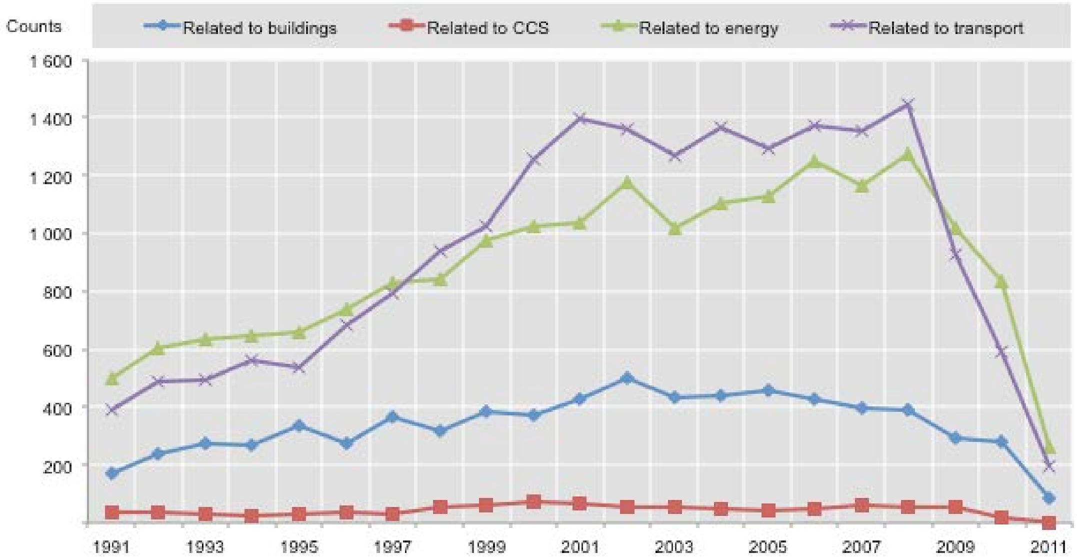
Y02B80/4	Y02T70/5
Y02B80/5	Y02T70/7
Y02B60/1	Y02T70/9
Y02B60/3	Y02T90/1
Y02B60/4	
Y02B60/5	
Y02B70/1	
Y02B70/3	
Y02B90/2	

Geothermal energy	Hydroelectricity	Nuclear fission	Nuclear fusion	Hydrogen	Fuel cells	Electric power conversion
Y02B10/4	Y02B10/5	Y02E30/3	Y02E30/1	Y02E60/3	Y02B90/1	Y02E60/6
Y02E10/1	Y02E10/2	Y02E30/4		Y02E70/1	Y02E60/5	
				Y02T90/4	Y02E70/2	
					Y02T90/3	

Electricity transmission and distribution	Energy storage	not classified
Y02E40/1	Y02E60/1	Y02B10/7
Y02E40/2	Y02E70/3	Y02E10/6
Y02E40/3	Y02E70/4	
Y02E40/4		
Y02E40/5		
Y02E40/6		
Y02E40/7		
Y02E60/7		

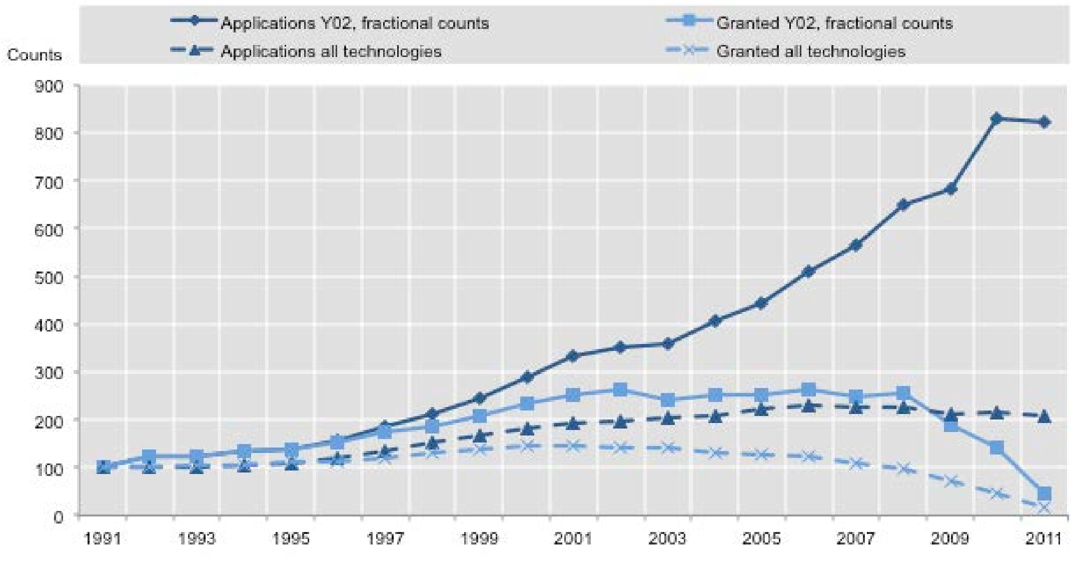
ANNEX C. ADDITIONAL FIGURES

Figure C.1: Granted patents by subcategories



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

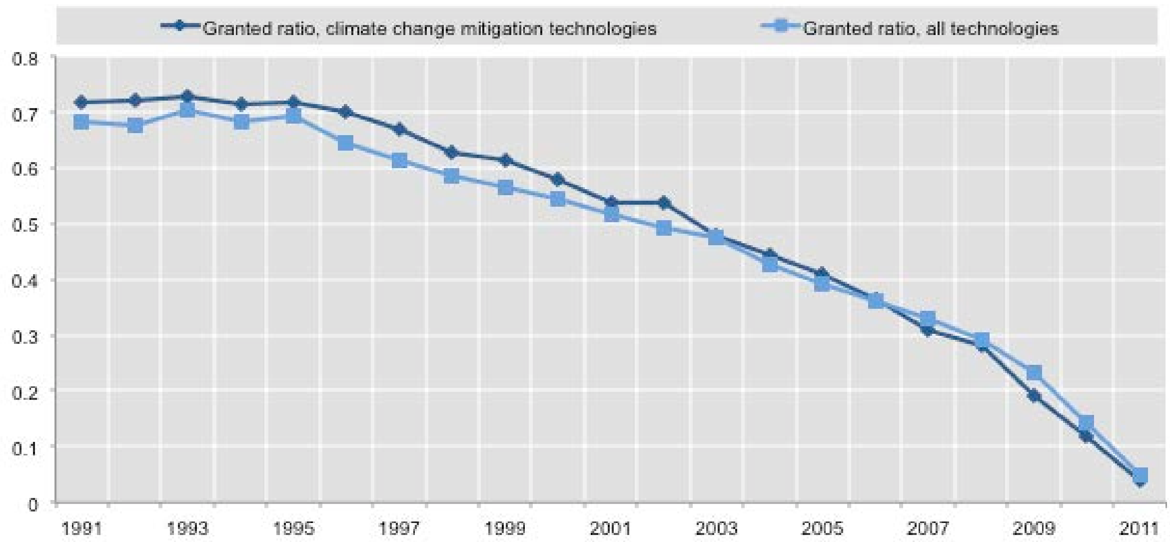
Figure C.2: Patent counts using fractional counts for Y02 technologies



Note: granted ratio is calculated as granted patents over total patent applications.

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

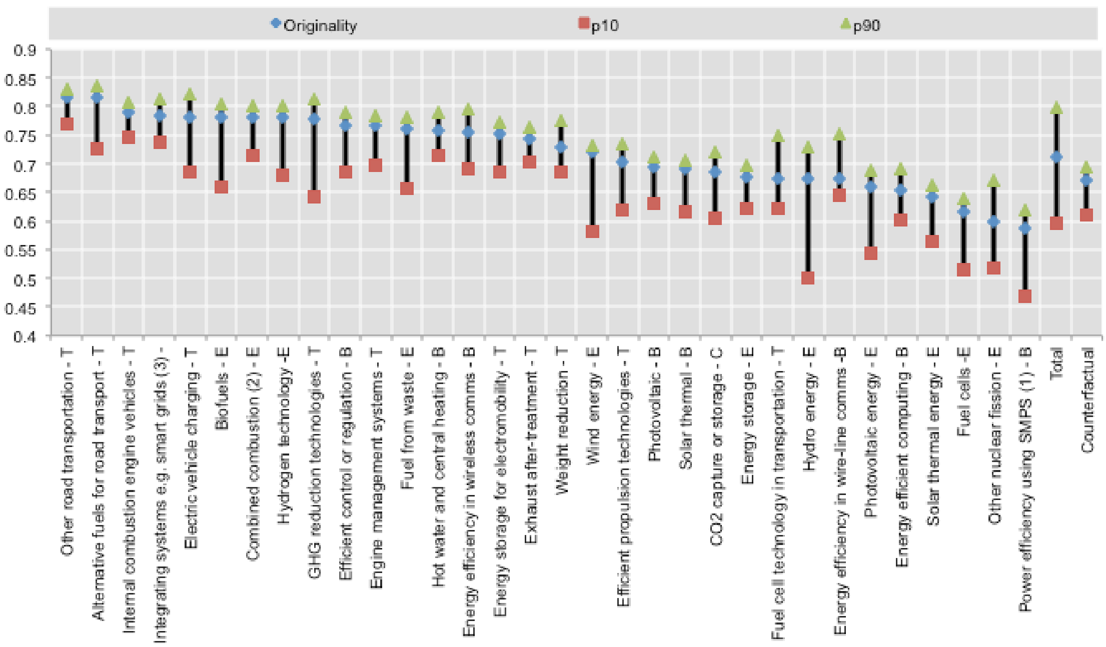
Figure C.3: Granted ratio for Y02 and counterfactual technologies



Note: granted ratio is calculated as granted patents over total patent applications.

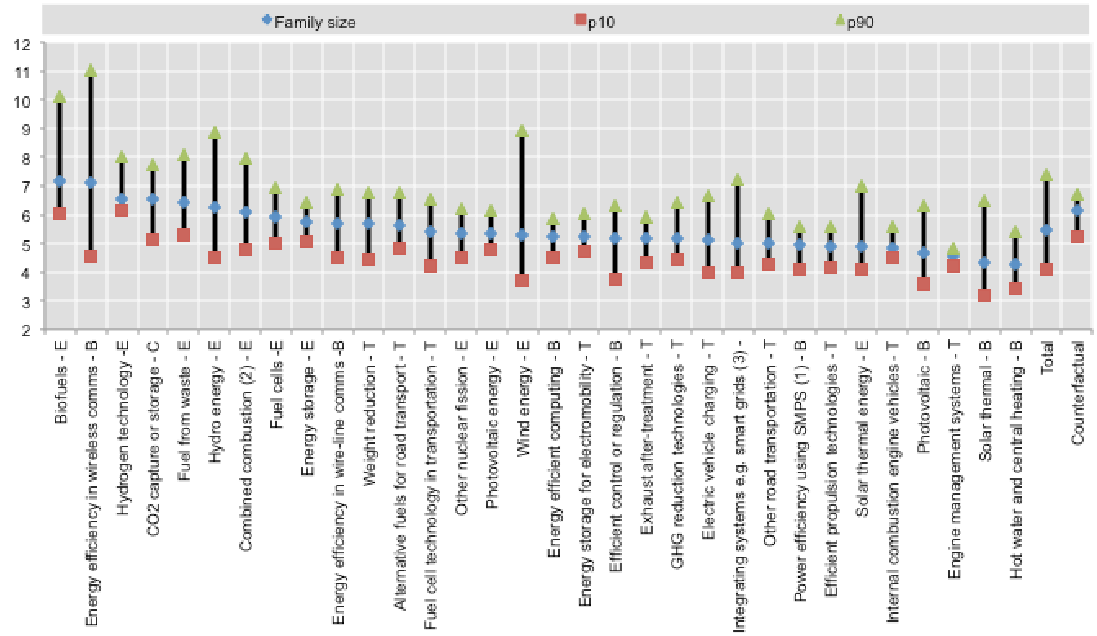
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Figure C.4: Originality 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



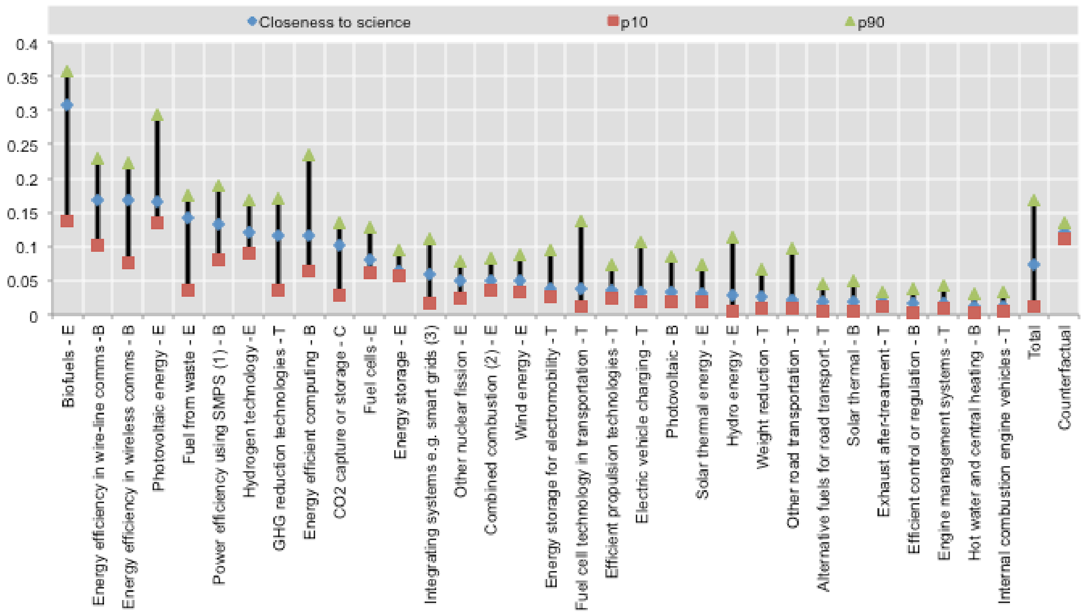
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Figure C.5: Family size 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



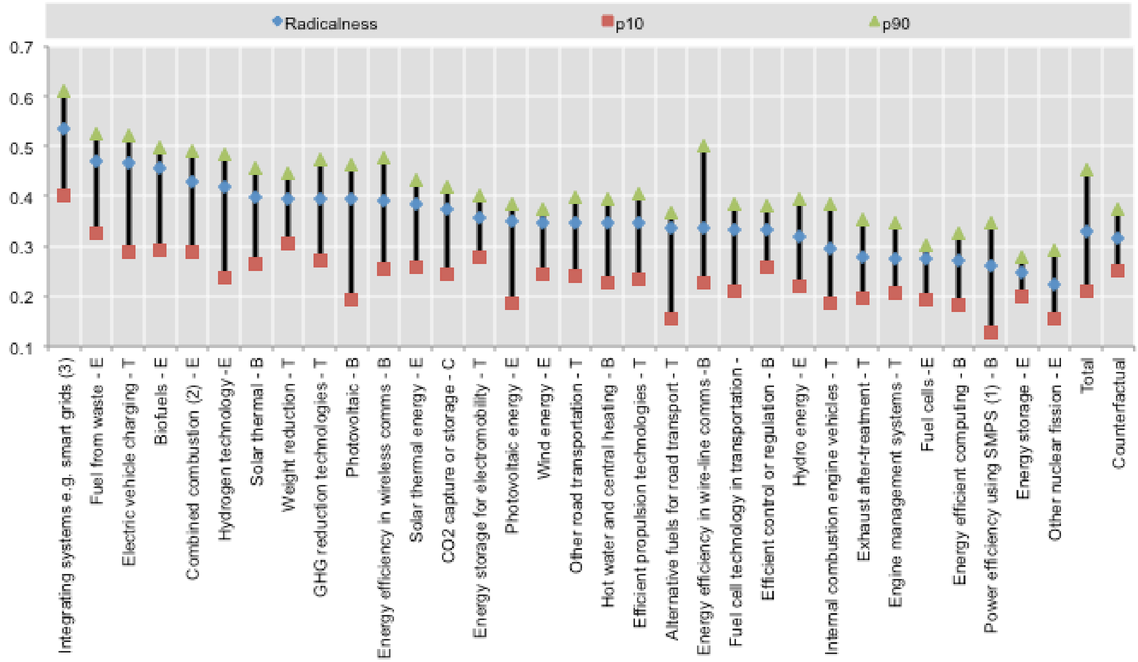
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Figure C.6: Closeness to science 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



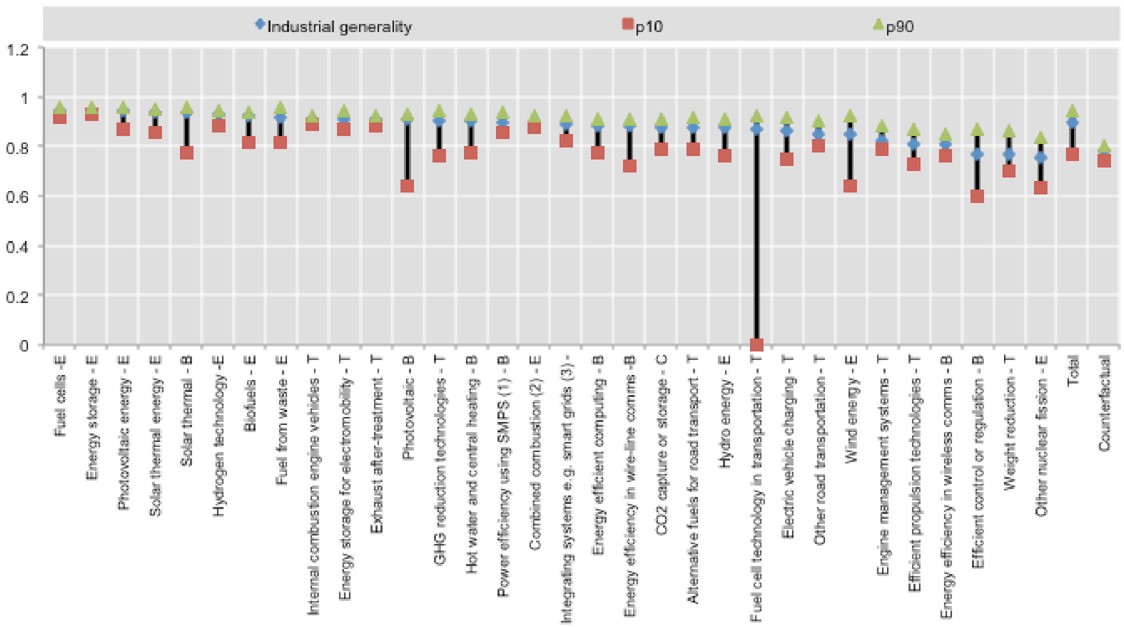
Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Figure C.7: Radicalness 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Figure C.8: Industrial generality 1991-2011 for all 7-digit CPC classes with >1000 patents (single counts)



Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

ANNEX D. RANKS OF TECHNOLOGY CLASSES

Table D.1. Quality indicator ranks of technology classes with >1000 patent applications all years, top 10 in red

CPC class	Label	Industrial Generality	Originality	Family size	Closeness to science	Radicalness	Patent applications
Y02B10/1	Photovoltaic [PV]	12	61	92	60	40	1322
Y02B10/2	Solar thermal	5	62	100	83	37	1842
Y02B30/1	using boilers	14	33	101	92	62	1375
Y02B30/7	Efficient control or regulation technologies	49	22	71	87	68	1111
Y02B60/1	Energy efficient computing	18	73	68	23	88	1781
Y02B60/3	Techniques for reducing energy-consumption in wire-line communication networks	19	68	42	11	66	1254
Y02B60/5	Techniques for reducing energy-consumption in wireless communication networks	38	34	7	13	41	3560
Y02B70/1	Technologies improving the efficiency by using switched-mode power supplies [SMPS], i.e. efficient power electronics conversion	15	90	80	18	90	1992
Y02B90/2	Systems integrating technologies related to power network operation and communication or information technologies mediating in the improvement of the carbon footprint of the management of residential or tertiary loads	17	13	78	40	3	1003
Y02C10/0	CO2 capture or storage	20	63	20	25	47	1676
Y02E10/2	Hydro energy	23	67	28	66	72	1139
Y02E10/4	Solar thermal energy	4	76	87	62	44	3401
Y02E10/5	Photovoltaic [PV] energy	3	72	60	14	58	10630
Y02E10/7	Wind energy	31	50	64	47	60	6168

Y02E20/1	Combined combustion	16	16	32	46	23	1987
Y02E30/4	Other aspects relating to nuclear fission	51	87	59	44	94	1033
Y02E50/1	Biofuels	7	15	5	3	18	3131
Y02E50/3	Fuel from waste	8	29	23	17	13	1644
Y02E60/1	Energy storage	2	65	40	33	91	17213
Y02E60/3	Hydrogen technology	6	17	19	21	25	2419
Y02E60/5	Fuel cells	1	79	38	28	87	10100
Y02T10/1	Internal combustion engine [ICE] based vehicles	9	10	88	93	81	8378
Y02T10/2	Exhaust after-treatment	11	42	72	85	85	5472
Y02T10/3	Use of alternative fuels	21	3	44	82	65	1036
Y02T10/4	Engine management systems	34	24	94	88	86	5481
Y02T10/6	Other road transportation technologies with climate change mitigation effect	30	2	79	78	61	4372
Y02T10/7	Energy storage for electromobility	10	38	69	54	53	6610
Y02T10/8	Technologies aiming to reduce greenhouse gasses emissions common to all road transportation technologies	13	18	73	22	39	1042
Y02T50/4	Weight reduction	50	47	43	69	38	1328
Y02T50/6	Efficient propulsion technologies	37	57	84	58	63	5046
Y02T90/1	Technologies related to electric vehicle charging	25	14	76	59	16	1881
Y02T90/3	Application of fuel cell technology to transportation	24	66	56	57	67	1013

Note: granted ratio is calculated as granted patents over total patent applications.

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

ANNEX E. ECONOMETRIC ANALYSIS: ROBUSTNESS CHECKS

Table E1 – Robustness: CPC group dummies

Dependent variable:	Patent granted 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Applicability 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	4.553***	4.563***	0.0573*	17.99***
	(0.346)	(0.329)	(0.0322)	(2.853)
Radicalness	-2.740***	-1.460*	-0.0763	25.73***
	(0.725)	(0.772)	(0.0766)	(9.652)
Closeness to science	0.881	2.512**	-0.269**	26.64**
	(1.040)	(1.160)	(0.128)	(11.45)
Family size	0.153*	0.134	0.00243	1.624***
	(0.0853)	(0.0870)	(0.00434)	(0.458)
Maturity	0.0528	0.0936	-0.0938***	-0.891
	(0.292)	(0.281)	(0.0211)	(2.825)
Patent apps. 1990-2000	0.462***	0.443***	0.0170***	2.112***
	(0.0751)	(0.0697)	(0.00380)	(0.474)
No. of backward cits.	0.0436	0.109	-0.0321***	1.291
	(0.110)	(0.114)	(0.0103)	(0.992)
Constant	0.572	0.983	1.470***	-23.83*
	(1.397)	(1.406)	(0.0941)	(14.31)
Observations	266	266	266	266
R-squared	0.839	0.813	0.573	0.461

*Note: All regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E2 – Robustness: one indicator at the time, patents granted

Dependent variable:	Patent granted 2001-2010 (logs)	Patent granted 2001-2010 (logs)	Patent granted 2001-2010 (logs)	Patent granted 2001-2010 (logs)	Patent granted 2001-2010 (logs)
Industrial generality	3.548***				
	(0.236)				
Originality		1.501**			
		(0.694)			
Radicalness			-1.471*		
			(0.770)		
Closeness to science				-1.063	
				(0.750)	
Family size					-0.0694
					(0.0481)
Observations	271	271	271	271	271
R-squared	0.682	0.328	0.326	0.321	0.322

*Note: All regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category and the constant term. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E3 – Robustness: one indicator at the time, patent applications

Dependent variable:	Patent applications 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Patent applications 2001-2010 (logs)
Industrial generality	3.453***				
	(0.227)				
Originality		1.517**			
		(0.707)			
Radicalness			-0.580		
			(0.798)		
Closeness to science				-0.296	
				(0.781)	
Family size					-0.0334
					(0.0515)
Observations	271	271	271	271	271
R-squared	0.661	0.312	0.301	0.300	0.301

*Note: All regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category and the constant term. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E4 – Robustness: one indicator at the time, applicability

Dependent variable:	Applicability	Applicability	Applicability	Applicability	Applicability
Industrial generality	0.101***				
	(0.0343)				
Originality		0.102			
		(0.0932)			
Radicalness			-0.0841		
			(0.0935)		
Closeness to science				-0.0692	
				(0.0843)	
Family size					-0.00947
					(0.00594)
Observations	271	271	271	271	271
R-squared	0.100	0.076	0.074	0.073	0.082

*Note: all regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category and the constant term. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E5 – Robustness: one indicator at the time, risk finance

Dependent variable:	Risk finance 2001-2010 (logs)	Risk finance 2001-2010 (logs)	Risk finance 2001-2010 (logs)	Risk finance 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	9.059***				
	(1.776)				
Originality		9.050**			
		(3.923)			
Radicalness			8.837**		
			(4.352)		
Closeness to science				1.367	
				(4.864)	
Family size					0.824**
					(0.325)
Observations	271	271	271	271	271
R-squared	0.164	0.101	0.099	0.087	0.115

*Note: all regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category and the constant term. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E6 – Robustness: controlling for R&D expenditures

Dependent variable:	Patent granted 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Applicability 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	4.834***	4.716***	0.0699**	18.57***
	(0.335)	(0.315)	(0.0318)	(2.844)
Radicalness	-3.530***	-2.369***	-0.0492	18.02*
	(0.922)	(0.854)	(0.0881)	(10.44)
Closeness to science	0.327	2.121**	-0.392***	28.61***
	(1.023)	(1.029)	(0.138)	(9.085)
Family size	-0.0221	0.0218	-0.0129**	1.528***
	(0.114)	(0.101)	(0.00637)	(0.504)
Maturity	0.0142	0.152	-0.143***	1.721
	(0.323)	(0.292)	(0.0217)	(2.357)
Patent apps. 1990-2000	0.389***	0.383***	0.00890**	2.028***
	(0.0846)	(0.0724)	(0.00383)	(0.421)
No. of backward cits.	0.215*	0.171	-0.00342	0.518
	(0.123)	(0.117)	(0.0107)	(1.068)
R&D expenditures	-9.84e-06	-2.20e-05***	2.51e-06	-0.000252***
	(8.91e-06)	(6.63e-06)	(1.52e-06)	(6.72e-05)
Constant	1.550	1.631	1.667***	-28.12**
	(1.558)	(1.377)	(0.0942)	(12.62)
Observations	268	268	268	268
R-squared	0.795	0.804	0.316	0.471

*Note: all regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.

Table E7 – Robustness: keeping only technologies with >20 patent applications in the 1990-2000 period

Dependent variable:	Patent granted 2001-2010 (logs)	Patent applications 2001-2010 (logs)	Applicability 2001-2010 (logs)	Risk finance 2001-2010 (logs)
Industrial generality	3.274***	3.179***	0.0621	12.03***
	(0.345)	(0.353)	(0.0436)	(2.748)
Radicalness	-2.091***	-0.978	-0.230**	13.51**
	(0.627)	(0.622)	(0.105)	(6.777)
Closeness to science	0.00518	0.902	-0.321**	2.392
	(0.843)	(0.941)	(0.161)	(9.064)
Family size	-0.00899	0.0746	-0.00973	2.221***
	(0.0578)	(0.0593)	(0.00758)	(0.503)
Maturity	0.0795	0.0853	-0.129***	-0.717
	(0.219)	(0.216)	(0.0319)	(2.050)
Patent apps. 1990-2000	1.028***	1.013***	0.00425	4.050***
	(0.375)	(0.354)	(0.00891)	(1.309)
No. of backward cits.	0.171	0.177	-0.0189	0.118
	(0.105)	(0.113)	(0.0132)	(0.860)
Observations	174	174	174	174
R-squared	0.655	0.617	0.354	0.239

*Note: all regression include CPC group dummies. All regressions are weighted by the no. of patents in the period 1990-2000 in the CPC category. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Source: OECD calculation based on EPO Worldwide Patent Statistical Database (PATSTAT) Spring 2014 and SDC.