Institute for
New Economic Thinking
AT THE OXFORD MARTIN SCHOOL

Knowledge Spillovers between Clean and Dirty Technologies

Su Jung Jee & Sugandha Srivastav

12th July 2022

INET Oxford Working Paper No. 2021-22

Knowledge Spillovers between Clean and Dirty Technologies

Su Jung Jee[1,2] and Sugandha Srivastav[1,3,4]

Does knowledge from dirty technologies spill over to clean technologies? The answer to this question has implications for the ease with which one can switch from dirty to clean R&D. Directed technical change models assume that there are no spillovers and consequently, find that clean technology subsidies are needed alongside carbon pricing. We empirically measure knowledge spillovers using data on patent citations from 1976-2020. The vast majority of clean technologies do not directly cite dirty technologies but are indirectly connected. Geothermal energy, clean metals and, carbon capture and storage have higher, but still modest, references to dirty knowledge.

Keywords: Knowledge spillovers, Directed Technical Change, Clean Technology, Dirty Technology

[1] *Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, UK.* [2] *University of Bradford, Faculty of Management, Law and Social Sciences, Richmond Road, Bradford BD7 1DP, UK.* [3]*Smith School of Enterprise and Environment, University of Oxford.* [4] *School of Geography and the Environment, University of Oxford *E-mails: s.j.jee@bradford.ac.uk; sugandha.srivastav@ouce.ox.ac.uk*

Introduction

The study of knowledge spillovers across firms, geographies and scientific fields helps scholars understand how information diffuses, where agglomeration occurs, and which areas of science have complementarities. This understanding is useful particularly if there is a need for *directed technical change.* While all innovation has value, some types of innovation can address pressing societal challenges. For example, the shift from capital-biased R&D to that which is labour-biased may help maintain social cohesion in the face of increasing levels of automation and joblessness. Similarly, the pivot from carbon-intensive ("dirty") innovation to low-carbon ("clean") innovation can help mitigate the impacts of climate change and air pollution. However, steering innovation in a particular direction requires understanding whether the two fields of knowledge have complementarities (even if their products are substitutes). Does dirty innovation spill over to clean innovation? Or is it largely distinct?

Governments around the world have adopted net-zero emissions targets to limit the impacts of anthropogenic climate change. Achieving these targets will require substantial amounts of clean innovation especially in sectors such as aviation, shipping, and iron & steel where low-carbon alternatives are limited/non-existent. However, the overall direction of innovation remains uncertain. Technological advances in hydraulic fracturing that occurred in the first decade of the new millennium have likely increased long-run emissions (Acemoglu et al. 2019) and, clean innovation, as measured via patents, peaked in 2010 and has subsequently fallen sharply (Fig. 1), alongside the drop in the number of clean tech start-ups (Popp et al 2020).

Models of directed technical change predict that due to path dependencies and switching costs, firms are more likely to continue innovating in dirty

technologies as they have done in the past, and that clean innovation subsidies will be needed alongside carbon pricing to achieve welfare improvements (Acemoglu et al. 2012, Acemoglu et al. 2016, and Aghion et al. 2016). However, this result depends on, among other things, the assumption that knowledge embedded in dirty technologies does not spill over into clean technologies.[1] If knowledge spills, then R&D switching costs may be overstated.

To conduct our analysis, we use citation data from patents granted by the U.S. Patent and Trademark office (USPTO) from 1976 to 2020, and leverage existing classifications of clean, grey and dirty patents in the literature.[2] We augment these classifications with some of our own manual tagging efforts (see Section III).[3] It is well understood that patents represent a subset of all innovation (OECD 2009), but since they are instruments to grant monopoly rights to novel inventions, they tend to capture frontier technologies with market potential, which is of particular interest since such technologies will be essential to achieve net-zero emissions. [4]

We measure direct and indirect knowledge spillovers between clean and dirty technologies using two metrics. First, by the proportion of backward citations in clean patents that directly refer to dirty patents (i.e., the "intensity metric"), and second, by the minimum distance between clean and dirty patents in the

[1] This is implicitly assumed by modelling scientific labor for dirty innovation as distinct from that for clean innovation.

[2] Grey patents are more efficient versions of dirty technologies.

[3] A patent fully discloses the invention such that a skilled practitioner can reproduce it (Perrons et al. 2021). There is thoroughness in ensuring relevant prior art is cited since patents are legal instruments that give monopoly rights and one must ensure the proposed invention is distinct from prior art.

[4] The American intellectual property regime is amongst the strongest in the world. For this reason, inventors from all over the world, particularly those whose inventions have high value, have sought protection in the US. In the literature, patents granted by the USPTO are often considered to represent the knowledge frontier (Granstrand, 2018).

citation network, (i.e., the "distance metric"). While the former metric has been used to measure knowledge spillovers frequently in the literature (e.g., Dechezleprêtre et al. 2014, Noailly and Shestalova 2017), the latter has not been used in this context and is complementary to the intensity metric.

The distance metric allows for the possibility that even if a clean patent does not directly cite a dirty patent, it may cite an intermediate technology which in turn cites a dirty patent, resulting in two (or more) degrees of separation.[5] The minimum distance finds the closest possible link between a clean and dirty technology in the citation network. For example, clean patent #36933371 which is about solar cells using high transmission glass, does not directly cite any dirty patent but it cites intermediate patent #22076423 which is about infrared absorbing glass, which in turn cites dirty patent #24632438, which is about inductively heating molten glass. This reflects an incremental and indirect way in which patented dirty knowledge has fed into a clean invention.

Patent citations are a reasonably good, albeit noisy, proxy knowledge spillovers. For example, Jaffe et al. (2000) survey R&D managers and find that citing inventors usually have close links or direct communication with cited inventors, and that reasons for patent citation include: using components of the cited technology and/or leveraging the cited technology to demonstrate the feasibility or use-case of the new invention. A more in-depth literature review of patent citation data's pros and cons can be found in Jaffe and De Rassenfosse (2019).

Our results show that less than one-tenth of clean technologies directly cite

[5] The distance metric is a measure of minimum distance between a clean and dirty patent in the citation network. One could also measure the average distance (i.e. calculate how far each reference is to a dirty patent, and take an average of the total) but this would result in a significant loss of information because a non-negligible proportion of references in clean patents are unconnected to dirty patents at any distance (i.e., having infinite distance).

dirty prior art. The mode of clean patents' distance to a dirty patent is 3 and over 53% of clean patents are within 3 degrees of separation from dirty patents. In terms of sectoral results, clean patents in geothermal energy, clean metals, carbon capture and storage (CCS), and long-haul transportation have bibliographies where around 6% of references relate to dirty technologies. For other clean sectors, the figure is significantly smaller. This shows that on the whole, there are very limited knowledge spillovers from dirty to clean technologies.

In terms of our contributions, literature to date has focused on whether clean or dirty technologies have higher spillovers (e.g., Dechezleprêtre et al. 2014, Noailly and Shestalova 2017) but this is different from asking how much clean technology learns from dirty. Understanding the degree of knowledge spillovers across these two fields can help policymakers better calibrate R&D switching costs and design innovation policy to achieve net-zero emissions.[6] Additionally, methodologically, we investigate indirect knowledge spillovers by measuring degrees of separation in the patent citation network. This is a novel way of thinking about indirect knowledge spillovers.

II. Trends in Clean and Dirty Innovation

Patenting related to clean technologies experienced a significant acceleration from the late 1990s and early 2000s. Reduced form evidence shows that rising energy prices contributed to this (e.g., Newell et al. 1999, Popp 2002, Verdolini and Gaelotti 2011). From the 1970s to the present day, the composition of clean patenting has changed: the share of patenting in electric and hybrid vehicles,

[6] We assume that all else equal, higher knowledge spillovers correspond to lower R&D switching costs.

clean ICT and solar PV has risen while it has declined in nuclear energy (Supplementary Fig. 6)

Clean patenting peaked in 2010 and has subsequently fallen due to the burst of the clean tech bubble, challenges in venture capital as a source of finance, and the rise of hydraulic fracturing (Dechezleprêtre 2017, Gaddy et al. 2017, Popp et al. 2020). This decline in clean patenting has raised concerns that current progress is not compatible with the aim to reach net-zero emissions by 2050.

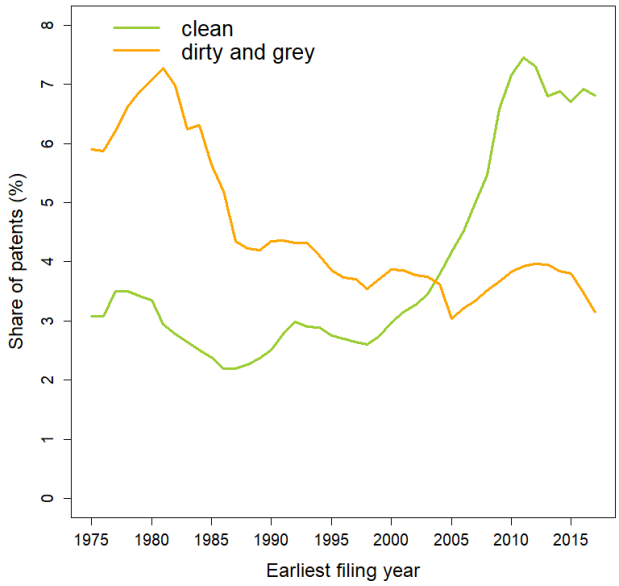


Figure 1. Number of patents as a share of all patents
Notes. The number of clean (green graph) and sum of dirty and grey (orange graph) patents as a share of all US patents by year. Patents are sorted by priority year (i.e., earliest filing year) and counted at DOCDB family-level.

The rise of dirty patenting in 2005, after years of decline was due to innovation in horizontal drilling and fracking (Supplementary Fig. 2b), which resulted in fossil fuel prices falling so much that natural gas became the primary fuel for electricity generation in the USA (Popp et al. 2020).

Path dependencies and technological lock-ins are cited to be an important barrier in advancing clean innovation at the required speed and scale (Acemoglu et al. 2012, Acemoglu et al. 2016, and Aghion et al. 2016). However, the extent to which knowledge from the dirty paradigm can or cannot be transferred to the clean paradigm is an empirical question which this paper will explore.

III. Data and Methods

A. Classification of Patents

We use World Patent Statistical Database (PATSTAT) (Spring 2021 version) to obtain the clean, dirty and grey patents issued by USPTO from 1976 to 2020 and the full patent citation network within the period, which consists of 55,188,499 links connecting patents at the family level. We define clean technologies as those that mitigate greenhouse gas emissions, dirty technologies as those that contribute to emissions, and grey technologies as energy-efficient versions of dirty technologies.[7] As an indicative illustration: renewable energy and electric vehicle patents are classified as clean, oil and gas patents are dirty, and energy-efficient methods of making steel are grey.

The relevant literature that we lean on to identify clean, dirty and grey patents includes: Haščič and Migotto (2015), Aghion et al. (2016), Popp et al. (2020) and Dechezleprêtre et al. (2021). Further, work by International Energy Agency provides a search strategy based on both cooperative patent classification (CPC) and keywords to find dirty patents associated with the fossil fuel supply chain (IEA 2021). Finally, while Aghion et al. (2016) identify grey patents for transport, we do not have such information for other sectors.

[7] For example, a more efficient internal combustion engine is an energy-efficient version of dirty technology. See Supplementary Table 1 for an explanation of what is classified as grey.

Therefore, we manually scan CPC codes for clean manufacturing and processing to identify additional grey patents (Supplementary Table 1).

The search strategy gives 258,078 clean patents, 145,753 dirty patents and 98,224 grey patents which are counted at DOCDB family level to avoid redundant counting of the same inventions. While any identification system is inevitably imperfect, our strategy represents a collection of all the existing clean, dirty and grey classifications alongside manual tagging, thereby improving upon what is standard practice in this literature. Finally, there are 31,053 patents that are classified as both clean and dirty, which we exclude from our baseline analysis to focus on the cases where the technologies are obviously clean or dirty (see Supplementary Fig. 1).

B. Methodology to Measure Spillovers

We develop two metrics to measure knowledge spillovers between clean and dirty technologies. The first is an "intensity metric" which checks the proportion of citations in a clean patent that refer to prior dirty patents.

The second is the distance metric, D_i, which calculates a clean patent i's minimum citation distance to prior dirty patent within the patent citation network (see Fig. 2). $D_i = 1$ represents clean patents that directly cite dirty patents, while $D_i \geq 2$ indicates clean patents that indirectly relate to prior dirty patents. That is, if $D_i \geq n$ (where n is larger than 2), a clean patent i must pass through at least n-1 non-dirty patents within the citation network to reach prior dirty patent(s). Clean patents that cannot be connected to dirty patents at any distance are treated as "unconnected."[8]

[8] The D metric is adapted to conduct additional analyses reported in supplementary material (e.g., Supplementary Fig. 3 shows connectivity from dirty to prior clean patents).

Our distance metric is adapted from work by Ahmadpoor and Jones (2017) who use it to measure the degrees of separation between scientific papers and patents. In our context, the interpretation we lend to D_i is as follows: as D_i increases, the likelihood that a dirty incumbent can leverage any of their existing R&D capabilities, either directly or indirectly, to diversify into a clean technology becomes smaller, all else equal (Cohen and Levinthal 1990, Lane and Lubatkin 1998). Another interpretation of D_i, which requires fewer assumptions on the ability of citations to trace knowledge flows, is that D_i measures proximity of fields of knowledge (i.e. greater values of D_i indicate lower levels of relatedness between those two fields).

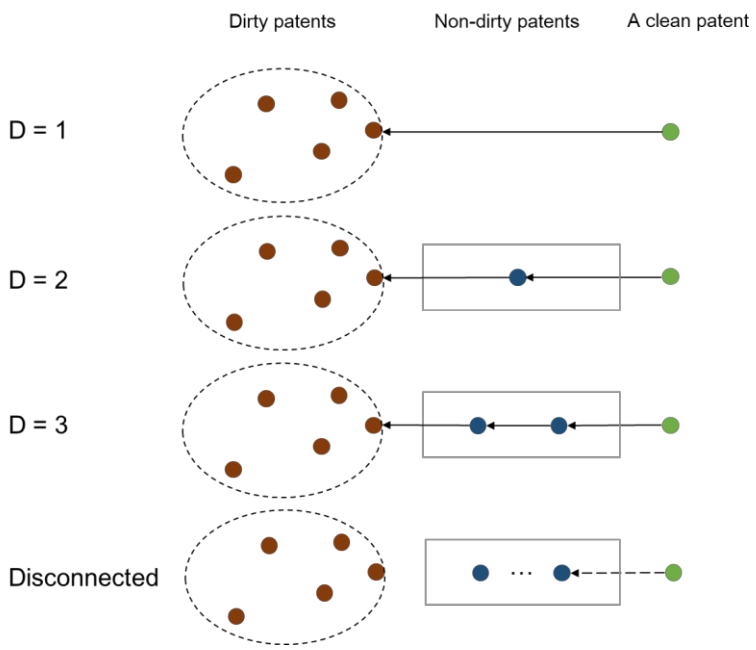


Figure 2. A schematic diagram showing the intellectual distances

Notes. This figure shows how the distance from a clean patent to prior dirty patent is determined. Arrows indicate the citation direction (from citing to cited patents).[9]

C. Discussion of Data

Notwithstanding various limitations including an incomplete coverage of inventive activity and a bias towards frontier technologies (e.g., OECD 2009), patents are used extensively in the study of innovation since they systematically document the inventor, owner of invention, field of invention, date of invention, and a list of prior relevant work (e.g., Griliches 1990, Jaffe et al. 1993).

Patent data is frequently used in the study of clean and dirty innovation thanks to the introduction of classification codes that help identify technologies that reduce or contribute to greenhouse gas emissions (e.g., Dechezleprêtre et al. 2013, Noailly and Shestalova 2017, and Perrons et al. 2021). Patents are also used in the literature to monitor how clean innovation responds to changes in prices and policies (e.g. Popp 2006; Johnstone et al. 2010; Aghion et al. 2016; Amore and Bennedsen 2016; Calel and Dechezleprêtre 2016, Popp 2020).

Patents have to cite prior work to demonstrate novelty. Citations can be added by the inventor, patent examiner and/or the patent attorney (Meyers 2000). While some citations are added by attorneys, the majority (62% in our sample) are added by the inventor who has a legal obligation to report all prior art which he or she consulted in the development of the invention. Citations added by the patent attorney or examiner are related to patented technology but may reflect inventions that the inventor was unaware of at the time of patenting (Berchicci and van de Vrande 2019). These citations help demonstrate that the patented technology took stock of all relevant prior knowledge and still produced a novel invention. For our baseline results we consider all citations regardless of who added them as we are broadly interested in knowledge spillovers across technological domains, rather than specifically, the inventor's knowledge base.

Real-life examples of how patent citations can trace knowledge spillovers are provided in Supplementary Materials, where we collect the backward citations from geothermal and CCS patents and show which dirty technologies they build upon (for example, we find geothermal inventions build upon prior dirty inventions related to underground power plants and heat exchangers). We also provide real-life examples of indirect knowledge spillovers at two degrees of separation: for example, clean patent #50731728 which is about reducing indoor air pollution, does not directly cite a dirty patent but cites intermediate patent #36590829 which is about removing contaminants, which in turn cites dirty patent #23777699 which is about particulate traps used in the exhaust system of a diesel engine.

IV. Results

A. Distance in the Citation Network

Only 7.5% of clean patents directly cite prior dirty patents.[10] The mode of clean patents' distance to a dirty patent is 3 and over 53% of clean patents are within three degrees of separation from dirty patents. (Fig. 3). By comparison, the average distance between *any* two randomly selected patents in USPTO is 8.5 (Mostafavi et al. 2012). The relative proximity between clean and dirty patents is partly attributable to the fact that many have common goals (e.g., generating electricity, fostering mobility, etc.). Yet the limited proportion of direct connections highlights that they are different technologies with limited direct knowledge spillovers on aggregate.[11]

[10] We are mostly interested in the relationship from clean to prior dirty, as the energy transition is about moving in the direction of "clean". However, for the reader's interest, we also calculate that 10% of dirty patents cite prior clean technologies (see Supplementary Fig. 3).

[11] Patent bibliography lengths are growing over time (Supplementary Fig. 5), which could theoretically

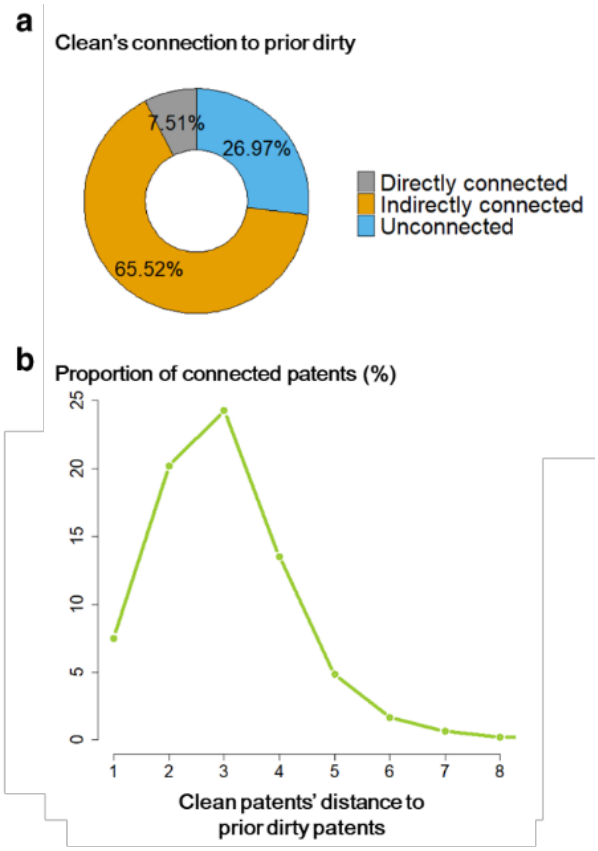


Figure 3. Connectivity from clean to prior dirty patents
Notes. Figures are made based on the directed graph of citation network from clean to previous dirty patents

B. Intensity of Connection

Taking all clean patents by sector, we check the proportion of references in their bibliographies that refer to prior dirty patents. The average clean patent has 9 references in its bibliography (at the DOCDB family-level) and 2% of backward citations that refer to dirty technologies.

Figure 4 presents a sectoral decomposition of the results. Geothermal energy, clean metals and CCS patents have the highest proportion of references that

increase the chance of citing old hydrocarbon knowledge and consequently, decrease D_i. However, this does not happen in our data (we see the opposite; values of D rise over time. See supplementary Fig. 6).

cite dirty inventions. Geothermal energy relies on geological surveying, drilling techniques, field development, and the construction of wells, pipelines, and other infrastructure, which requires knowledge inputs that are commonly used by fossil fuel firms. Clean innovation in metals is largely incremental in nature and consequently, still embedded in the dirty production paradigm.[12] Finally, CCS is a complement to coal-fired power plants, gas stations and other point-sources of carbon emissions and naturally has to be fitted to these.

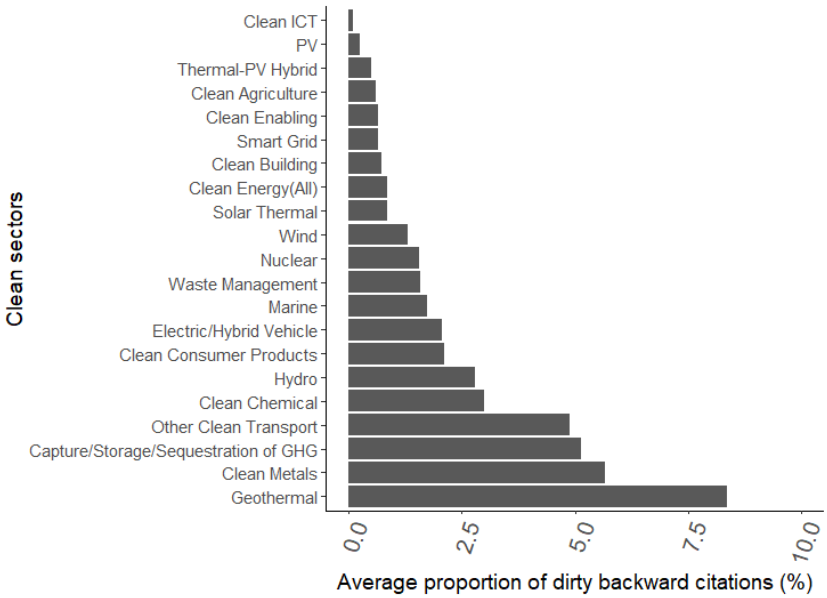


Figure 4. Sectoral details of intensity metric
Notes. Each clean sector's average proportion of dirty backward citations

Clean ICT and solar PV have negligible direct links to dirty technologies. In the case of solar PV, this may be reflective of just how different the photovoltaic paradigm is from the hydrocarbon paradigm. The former is based on the photovoltaic effect while the latter relies on spinning a coil around a magnet to

[12] There are some radical strands zero-carbon innovation in metals such as zero-carbon steel made from hydrogen (e.g., the HYBRIT project). Such innovation effort is still so nascent and rare that it is not reflected in patent databases. As such, it is possible that future analyses find that clean metals rely less on dirty knowledge because radical clean innovation becomes more commonplace and better documented.

generate power (i.e., turbines). This also explains intuitively why turbine-based technologies, such as hydroelectric power, wind energy, and some types of marine energy have more citations to dirty technologies than solar PV.

Marine energy, for example, requires knowledge inputs that are common to dirty technologies such as offshore oil. This includes seabed engineering, constructing offshore platforms, placing under-sea cables, under-sea robots and materials that can withstand biofouling. This perhaps explains why offshore oil companies like British Petroleum put in bids for seabed rights in the North Sea to develop offshore wind farms, as they can leverage existing knowhow (King 2021).

For electric/hybrid vehicles some elements of innovation such as car design and a more efficient internal combustion engine rely on dirty knowledge, while other elements, such as batteries are different.

C. Indirect Connections

The majority of the connections between clean and dirty technologies are indirect (see Fig. 3 and Fig. 5). The distance at which clean patents are connected to dirty patents differs largely by technology, as plotted in Fig. 5.

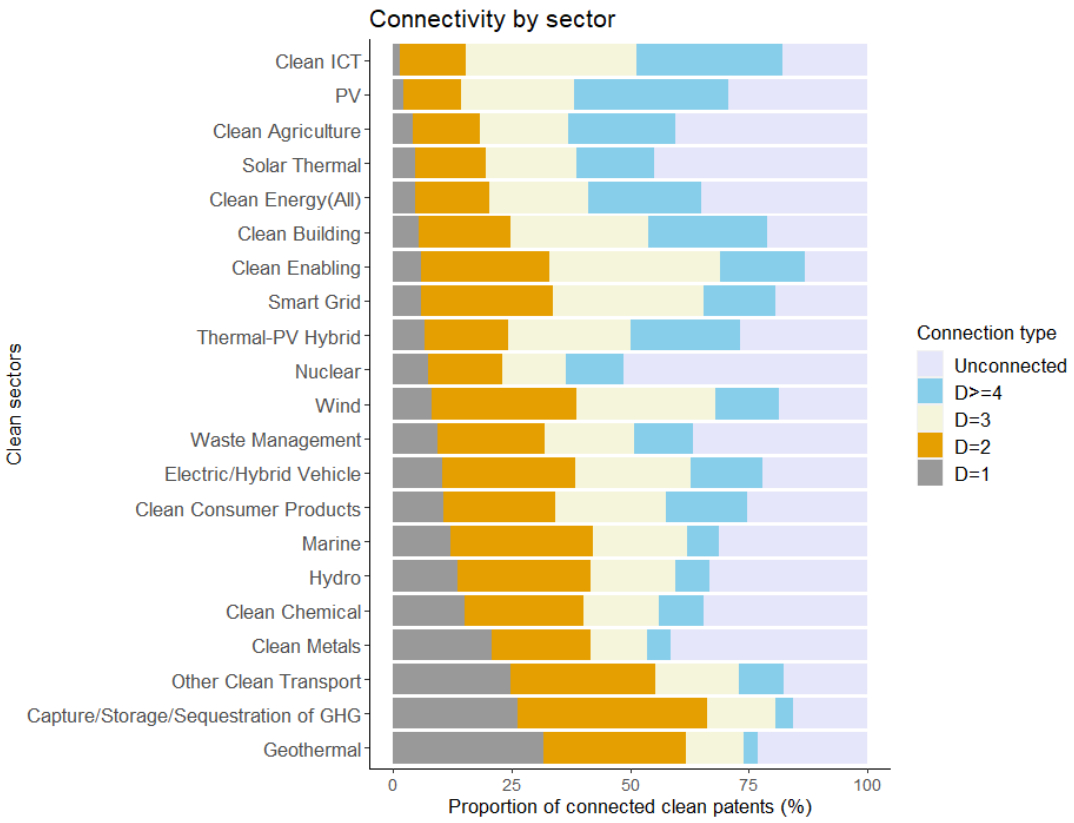


Figure 5. Sectoral details of direct and indirect connections
Notes. Each clean sector's direct and indirect connections to prior dirty technologies.

Electric/hybrid vehicles and wind power have a relatively low share of direct connections but a high share of indirect connections at distance 2 (i.e. they cite technologies that cite dirty technologies). Certain technologies such as nuclear energy have amongst the highest proportion of patents that are unconnected to dirty technologies in the citation network, illustrating how distant the two fields are.

Overall, these results imply dirty innovators are better placed to diversify into clean innovation related to geothermal energy, CCS, long-haul transportation ("other transport") and clean metals than clean technologies such as solar PV or clean ICT from a pure knowledge perspective. However, in practice, diversification requires considering not only knowledge spillovers but

also other variables such as production capabilities, access to raw materials, supply chains, etc. Nevertheless, since knowledge is a key input into the production process, understanding the spillovers between clean and dirty technologies is important. On an aggregate level, knowledge spillovers are limited, as 92.5% of clean patents have no direct citation to a dirty patent and a non-trivial proportion (27%) of clean patents are not connected to a dirty patent at any distance.

D. Grey Technologies

Grey technologies can either be a steppingstone towards decarbonisation or lock-in emissions. For example, in areas where clean energy is available, building new gas infrastructure may simply exacerbate carbon lock-in even though it is less polluting than coal. On the other hand, if no radically clean substitutes are available or if they are imperfect substitutes, grey technologies can help incrementally reduce emissions while more radical solutions are being developed (Stern and Valero 2021). We find that 23% of all grey patents have at least one reference to dirty patents (D=1). This is substantially higher than the 7.5% of clean patents that have at least one reference to dirty patents. Additionally, on average, 7% of all backward citations in grey patents relate to dirty patents. This indicates, perhaps intuitively, that pivoting from dirty to grey R&D is relatively more accessible than dirty to clean R&D due to the higher degree of knowledge spillovers.

E. Cross Technology Linkages

Which dirty technologies do clean technologies draw on? Supplementary Fig. 8 maps clean technologies to connected dirty technologies in a disaggregated manner. Electric vehicles, smart grids, enabling technologies in manufacturing,

clean buildings, and clean ICT are connected to dirty transportation but at increasing levels of indirectness (see star dots in Supplementary Fig. 8a). Clean technologies in hard-to-decarbonise sectors, such as chemicals, metals and long-distance transportation share connections with dirty manufacturing and downstream fossil fuel technologies (see triangle and diamond dots in Supplementary Fig. 8a). Geothermal cites upstream fossil fuels likely due to the common reliance on drilling and geological surveying (see yellow dots in Supplementary Fig. 8b).

Firms that specialize in innovating in mining, drilling, processing may be in a relatively better position to pivot into clean innovation which also rely on some of these engineering processes. This may provide a way of partially ameliorating the transitional costs since some proportion of existing knowhow will find gainful employment in "adjacent" clean sectors (e.g. Hausmann et al. 2014)

V. Discussion

As the fracking revolution revealed, the overall direction of innovation is still uncertain and in the absence of policy, it can move in directions that do not support the energy transition. The target of net-zero emissions by 2050 necessitates control over the direction of innovation, and in this regard, it is important to understand areas where dirty knowledge spills over into clean technologies.

The seminal work on directed technical change and the environment argues that due to the market size effect, carbon prices and targeted clean innovation subsidies are needed to pivot innovation towards technologies that reduce emissions and increase overall welfare (Acemoglu et al. 2012). This model implicitly assumes that spillovers between clean and dirty technologies are non-

existent. We validate this assumption empirically by showing that the vast majority of clean technologies have no direct links to dirty technologies; and that even for clean technologies that do cite dirty patents, the share of dirty patents in the bibliography is limited. We also measure indirect citations and find that 27% of clean patents are not connected to a dirty one at any distance in the citation network, and that over time, the share of clean patenting in sectors that are distant, such as solar PV and clean ICT is growing. This points towards the idea that clean and dirty knowledge are largely distinct, and that switching costs are indeed likely to be non-trivial, underscoring the importance of directed clean technology subsidies.

Aggregate results do, however, conceal areas of some spillovers. In reality, the degree to which clean technology learns from dirty technology is nuanced, sector-specific and sometimes indirect. Clean sectors that draw from the hydrocarbon knowledge paradigm include geothermal, clean metals, CCS, long-haul transport and clean chemicals. Others include marine and wind energy where the learning largely is indirect rather than direct. These areas of relative proximity may offer dirty R&D incumbents some opportunities to recycle their knowhow.

In terms of limitations, we are constrained, to a large extent, by existing classifications of clean and dirty patents. Future work could use machine learning to better discern categories using information contained within patents' abstract, title and claims. Additionally, while patents represent a well-codified and accessible dataset for researchers, not all forms of technological knowledge are captured in patents. Some knowledge remains undisclosed, such as trade secrets and tacit knowledge of within-firm expertise (Roach and Wesley 2013).

An avenue for future research is to explore knowledge flows between clean technologies and other general sectors of the economy. Furthermore, one could use data on scientific publications, categorise them as clean, dirty or grey, and measure spillovers using this dataset, and compare it to the results of this paper.

Lastly, it would be interesting to more systematically test if knowledge spillovers, as measured by patent citations, are positively correlated with the use of similar capital and labour (e.g., seabed engineers, floating platforms, robots that work at sea, etc.).

VI References

Acemoglu, D., Aghion, P., Barrage, L. and Hemous, D., 2019. Climate Change, Directed Innovation, and Energy Transition: The Long-run Consequences of the Shale Gas Revolution. *Working Paper.*

Acemoglu, D., Aghion, P., Bursztyn, L. and Hemous, D., 2012. The Environment and Directed Technical Change. *American Economic Review*, *102*(1), Pp.131-66.

Acemoglu, D., Akcigit, U., Hanley, D. and Kerr, W., 2016. Transition to Clean Technology. *Journal of Political Economy*, *124*(1), pp.52-104.

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., and Van Reenen, J. 2016. Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from The Auto Industry. *Journal* of *Political Economy*, *124*(1), 1-51.

Ahmadpoor, M., and Jones, B. F. 2017. The Dual Frontier: Patented Inventions and Prior Scientific Advance. *Science*, *357*(6351), 583-587.

Amore, M. D. and M. Bennedsen 2016. Corporate governance and green innovation. *Journal of Environmental Economics and Management* 75 (C), 54.72.

Berchicci, L., and Van De Vrande, V. 2019. Noisy Or Valuable? The Effect of Examiner-Added Citations on Firm Knowledge Flows. In DRUID Conference.

Calel, R. and A. Dechezleprêtre 2016. Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market. *The Review of Economics and Statistics* 98 (1), 173.191.

Cohen, W. M., and Levinthal, D. A. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 128-152.

Dechezleprêtre, A., Martin, R., and Mohnen, M. 2014. Knowledge Spillovers from Clean and Dirty Technologies.

Dechezleprêtre, A. 2017. "Sustaining Investment in Climate Innovation." Climate Innovation Insights, Series 1.2: Accelerating the Evolution of Climate Innovation Clusters, Climate–KIC. Https://Www.Climate-Kic.Org/Wp Content/Uploads/2017/03/Insight02_Proof4.Pdf

Dechezleprêtre, A., Muckley, C. B., and Neelakantan, P. 2021. Is Firm-Level Clean or Dirty Innovation Valued More? *The European Journal* of *Finance*, *27*(1-2), 31-61.

Gaddy, B. E., Sivaram, V., Jones, T. B., and Wayman, L. 2017. Venture Capital and Cleantech: The Wrong Model for Energy Innovation. *Energy Policy*, *102*, 385-395.

Granstrand, O. 2018. Evolving Properties of Intellectual Capitalism: Patents and Innovations for Growth and Welfare, Edward Elgar Publishing, Cheltenham

Haščič, I., and Migotto, M. 2015. Measuring Environmental Innovation

Using Patent Data 89. Pp. 0–1 OECD Environment Working Papers.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., and Simoes, A. 2014. *The Atlas* of *Economic Complexity: Mapping Paths to Prosperity.* MIT Press.

IEA (International Energy Agency). 2021. Methodology for Identifying Fossil Fuel Supply Related Technologies in Patent Data.

Jaffe, A. B., Trajtenberg, M., Henderson, R. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3), 577–598.

Jaffe, A. B., De Rassenfosse, G., 2019. Patent citation data in social science research: Overview and best practices. *Research Handbook on the Economics of Intellectual Property Law.*

Johnstone, N., I. Hascic, and D. Popp 2010. Renewable Energy Policies and Technological Innovation: Evidence based on Patent Counts. *Environmental and Resource Economics* 45 (1), 133.155.

King, I. 2021. BP Blows Away Bid Rivals with Big Bet on Wind Farms. [Online] Sky News. Available At: Https://News.Sky.Com/Story/Bp-Blows-Away-Bid-Rivals-With-Big-Bet-On-Wind-Farms-12212580 [Accessed 13 Nov. 2021].

Lane, P. J., and Lubatkin, M. 1998. Relative Absorptive Capacity and Interorganizational Learning. *Strategic Management Journal*, *19*(5), 461-477.

Mostafavi, S., Goldenberg, A., and Morris, Q. 2012. Labeling Nodes Using Three Degrees of Propagation. *Plos One*, *7*(12), E51947.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. 1999. The Induced Innovation Hypothesis and Energy-Saving Technological Change. *The*

Quarterly Journal of *Economics*, *114*(3), 941-975.

Noailly, J., and Shestalova, V. 2017. Knowledge Spillovers from Renewable Energy Technologies: Lessons from Patent Citations. *Environmental Innovation* and *Societal Transitions*, *22*, 1-14.

OECD 2009. OECD Patent Statistics Manual. Report.

Popp, D., 2002. Induced Innovation and Energy Prices. *American Economic Review*, *92*(1), Pp.160-180.

Popp, D. 2006. International innovation and diffusion of air pollution control technologies: the effects of NOX and SO2 regulation in the US, Japan, and Germany. *Journal of Environmental Economics and Management 51 (1), 46.71.*

Popp, D., Pless, J., Haščič, I. and Johnstone, N. 2020. Innovation and Entrepreneurship in The Energy Sector (No. W27145). *National Bureau* of *Economic Research.*

Stern, N., and Valero, A., 2021. Research Policy, Chris Freeman Special Issue Innovation, Growth and The Transition to Net-Zero Emissions. *Research Policy*, *50*(9), 104293.

Roach, M., and Wesley M C. 2013. "Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research." *Management science* vol. 59,2 (): 504-525.

Verdolini, E., and Galeotti, M. 2011. At Home and Abroad: An Empirical Analysis of Innovation and Diffusion in Energy Technologies. *Journal* of *Environmental Economics* and *Management*, *61*(2), 119-134.

Author Contributions

Both authors contributed equally to the manuscript.

Acknowledgements

Authors would like to thank Alex Teytelboym, Richard Green, Jacquelyn Pless, Cameron Hepburn, Pia Andres, Sam Fankhauser, François Lafond, Eugenie Dugoua, Tim Dobermann, Joris Bücker, Kerstin Hötte, Brian O'Callaghan, Verena Weidmann and INET seminar participants for their feedback. Su Jung Jee acknowledges support from Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1A6A3A03037237). Sugandha Srivastav acknowledges support from the Oxford Martin School Programme on the Post Carbon Transition, the Institute of New Economic Thinking, as well as the Climate Compatible Growth programme.

VII Appendix

Supplementary Table 1. Patent search strategy

	Sector	Sub-sector	CPC	IPC
Clean[1]	Clean Building		Y02B	
	Capture/Storage/Sequestration of GHG (CCS)		Y02C	
	Clean Energy	Geothermal	Y02E 10/10	
		Hydro	Y02E 10/20	
		Nuclear	Y02E 30	
		Photovoltaic	Y02E 10/50	
		Sea	Y02E 10/30	
		Solar thermal	Y02E 10/40	
		Thermal-PV hybrids	Y02E 10/60	
		Wind	Y02E 10/70	
		Others	Y02E 40, Y02E 50, Y02E 60, Y02E 70	
	Clean ICT		Y02D	
	Clean Manufacturing[2]	Agriculture	Y02P 60	
		Chemical	Y02P 20, Y02P 30, Y02P 40	
		Consumer products	Y02P 70, Y02P 80	
		Enabling	Y02P 90	
		Metal	Y02P 10	
	Smart Grid		Y04S	
	Clean Transport[3]	Electric/Hybrid vehicle	Y02T 10, Y02T 90	see Aghion et al. (2016)[4]
		Aero, waterways, or railways	Y02T 30, Y02T 50, Y02T 70	
	Waste Management		Y02W	
Dirty	Energy	Upstream	see IEA (2021)[5]	
		Processing and downstream		
		Transmission and distribution		
	Transport	Internal combustion engine		see Aghion et al. (2016)
	General (Manufacturing)			see Dechezleprêtre et al. (2021)

	Energy[6]	Combustion technologies with mitigation potential	Y02E 20	
		Technologies related to metal processing/Recycling	Y02P 10/20	
		Technologies related to metal processing/Process efficiency	Y02P 10/25	
		Technologies relating to chemical industry/Process efficiency/Energy recovery, e.g., by cogeneration, H2recovery or pressure recovery turbines	Y02P 20/129	
		Technologies relating to chemical industry/Feedstock/the feedstock being recycled material, e.g., plastics	Y02P 20/143	
		Technologies relating to chemical industry/Reduction of greenhouse gas [GHG] emissions, e.g., CO2	Y02P 20/151/low	
		Technologies relating to chemical industry/Improvements relating to chlorine production	Y02P 20/20/low	
		Technologies relating to chemical industry/Improvements relating to adipic acid or caprolactam production	Y02P 20/30/low	
		Technologies relating to chemical industry/Improvements relating to fluorochloro hydrocarbon, e.g., chlorodifluoromethane [HCFC-22] production	Y02P 20/40/low	
		Technologies relating to chemical industry/Improvements relating to the production of bulk chemicals	Y02P 20/50/low	
Grey[8]	Manufacturing[7]	Technologies relating to oil refining and petrochemical industry/Ethylene production	Y02P 30/40	
		Technologies relating to the processing of minerals/Production of cement, e.g., improving or optimising the production methods; Cement grinding/Energy efficiency measures, e.g., improving or optimising the production methods	Y02P 40/121	
		Technologies relating to the processing of minerals/Glass production, e.g., reusing waste heat during processing or shaping/Improving the yield, e-g-reduction of reject rates	Y02P 40/50, Y02P 40/57	
		Technologies relating to the processing of minerals/Production of ceramic materials or ceramic elements, e.g., substitution of clay or shale by alternative raw materials, e.g., ashes	Y02P 40/60	
		Climate change mitigation technologies in the production process for final industrial or consumer products/Manufacturing or production processes characterised by the final manufactured product	Y02P 70/50/low	
		Climate change mitigation technologies for sector-wide applications/Reducing waste in manufacturing processes; Calculations of released waste quantities	Y02P 80/30	
		Climate change mitigation technologies for sector-wide applications/Minimising material used in manufacturing processes	Y02P 80/40	
	Transport	Fuel efficiency of internal combustion engine-based vehicles	Y02T 10/10	see Aghion et al. (2016)

Notes. 1. We refer to Haščič and Migotto (2015), Popp et al. (2020) and Aghion et al. (2016) to collect clean patents; 2. Grey manufacturing patents are excluded from the list of clean manufacturing patents; 3. Grey transport patents are excluded from the list of clean transport patents; 4. B06L 11 is transferred to B60L 50/00 - B60L 58/40 (see IPC ver. 2019); 5. Search strategy

of IEA (2021) combines both CPC and keyword based search; 6. Authors' interpretation; 7. Authors' interpretation; 8. Grey patents are excluded from the final list of clean and dirty patents. Patents classified as both clean and dirty patents are excluded from our main analysis (see Supplementary Fig. 1).

Supplementary Table 2. Top 10 countries of creating clean, dirty and grey patents

	Clean patents		Dirty patents		Grey patents	
	Country	Patent counts (Family)	Country	Patent counts (Family)	Country	Patent counts (Family)
1	US	122,515	US	88,319	US	39,660
2	Japan	56,632	Japan	19,377	Japan	28,462
3	Germany	22,622	Germany	13,344	Germany	13,930
4	South Korea	18,755	UK	6,546	South Korea	3,530
5	France	9,230	France	6,465	France	3,422
6	China	8,172	Canada	5,675	UK	3,394
7	Taiwan	7,312	Switzerland	2,177	Canada	2,084
8	UK	6,914	Netherlands	2,081	Italy	1,704
9	Canada	6,472	Sweden	1,669	Switzerland	1,477
10	Switzerland	3,327	Norway	1,529	China	1,244

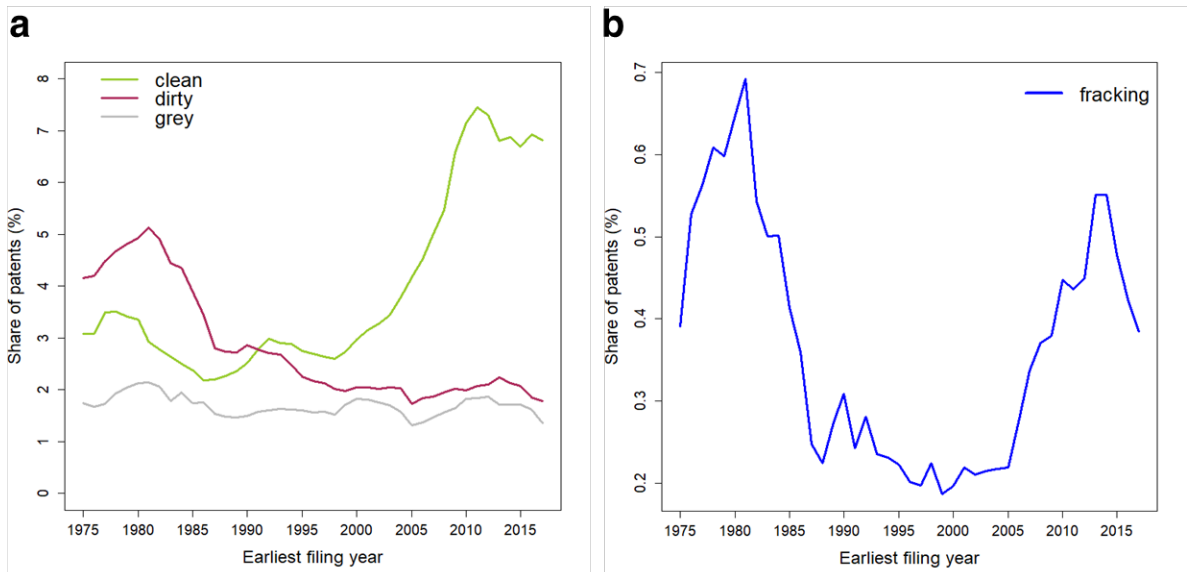
Notes. Patents issued by USPTO during 1976-2020

Supplementary Figures

Supplementary Figure 1. Venn diagram for clean, dirty and grey patents

Notes. We regard clean or dirty patents that are also classified grey patents as grey patents. Patents classified as both clean and dirty patents are excluded from our main analysis.

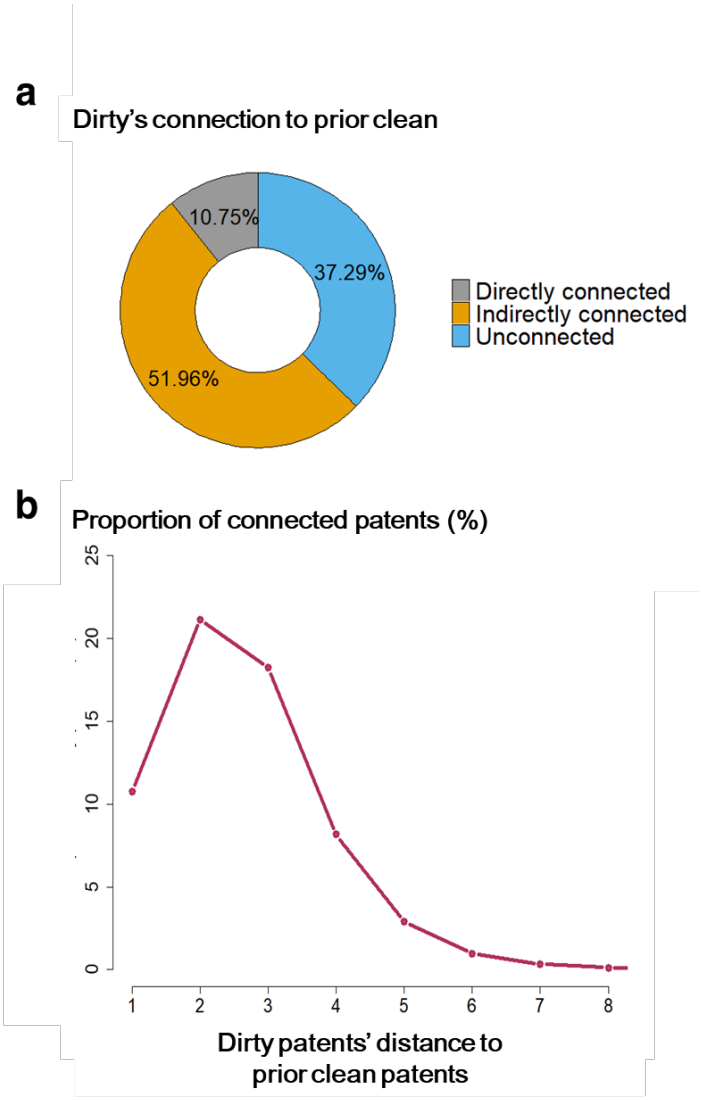
Supplementary Figure 2. Number of patents as a share of all patents



Notes. (a) The number of clean, dirty and grey patents as a share of all the US patents by year, (b) The number of fracking related patents as a share of all US patents by year. To search fracking related patents, we combine the search strategy for unconventional fossil fuel patents in IEA (2021) and for hydrofracturing patents in Popp et al. (2020). Patents are sorted by priority year (i.e.,

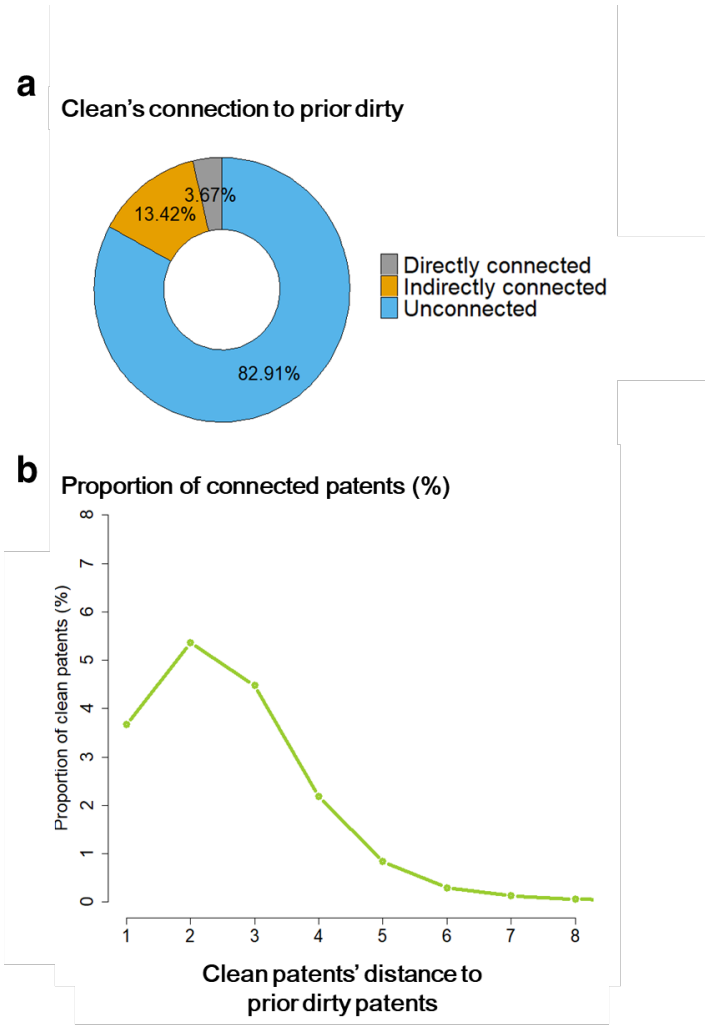
earliest filing year) and counted at DOCDB family-level.

Supplementary Figure 3. Connectivity from dirty to prior clean patents



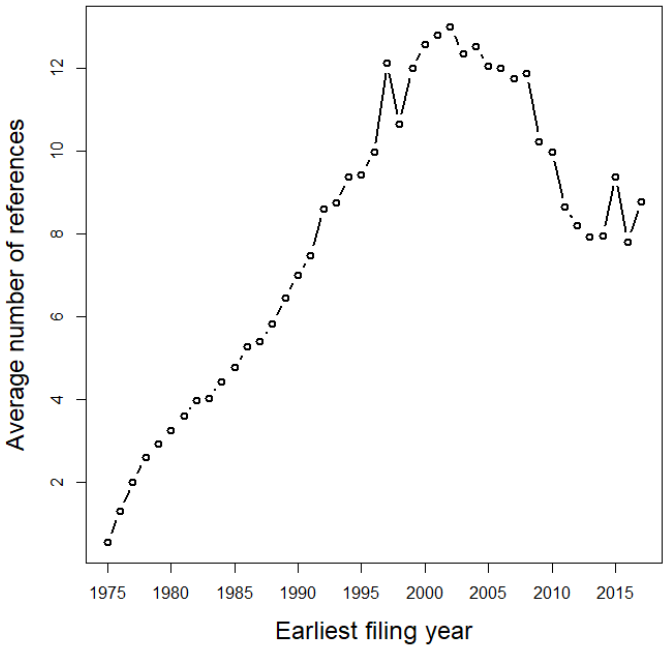
Notes. Figures are made based on the directed graph of citation network from dirty to previous clean patents.

Supplementary Figure 4. Connectivity from clean to prior dirty patents (Applicant citation only)

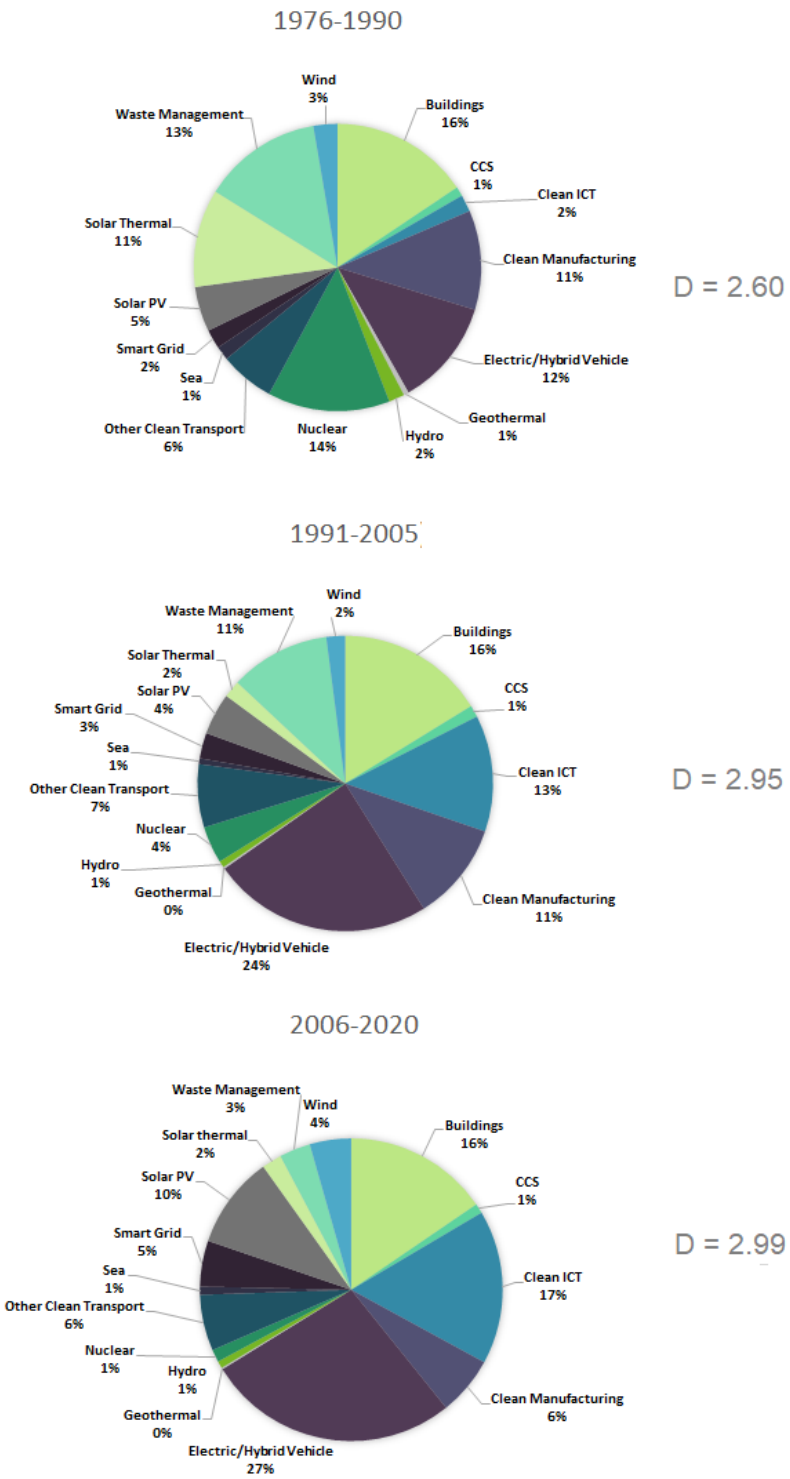


Notes. Figures are made based on the directed graph of citation network from clean to previous dirty patents.

Supplementary Figure 5. Bibliography length of clean patents



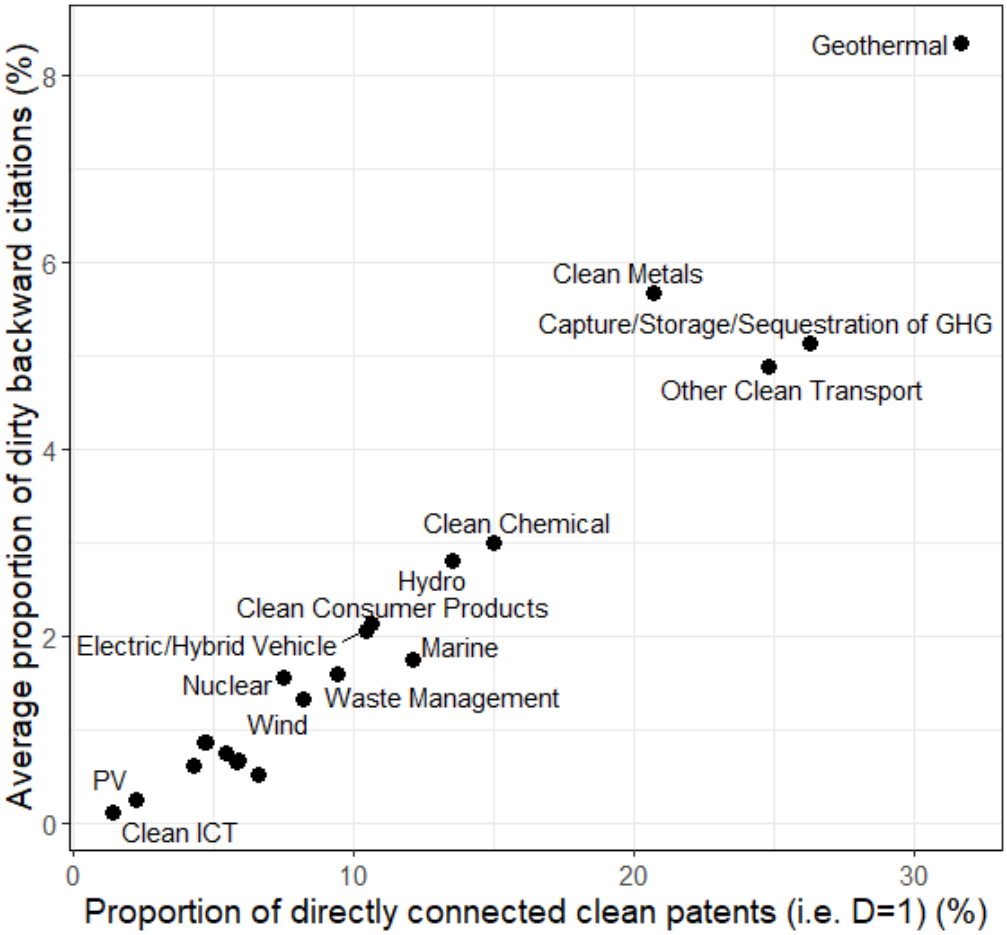
Supplementary Figure 6. The Composition of Clean Patenting over Time



Notes. Graph shows share of granted patents in each time period. Total count for first, second and third

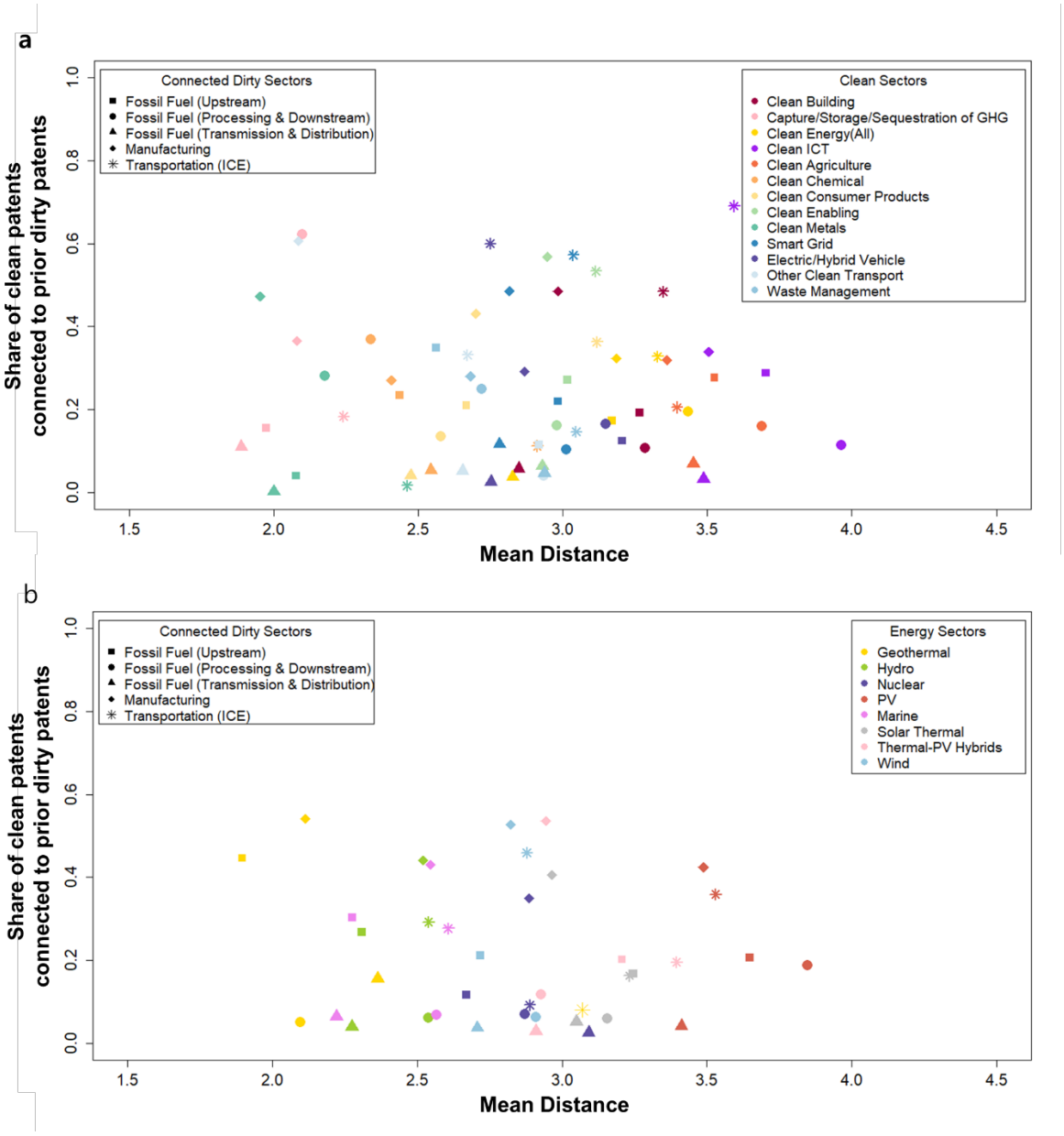
time period is respectively, 28,000, 60,000, and 140,000 (reported to two significant figures). Patents can be tagged as relevant to multiple clean sectors.

Supplementary Figure 7. Sectoral mapping of range and intensity of direct connections



Notes. X-axis indicates the proportion of clean patents that are directly connected to prior dirty patents among the clean patents in each sector (Range of direct connection). Y-axis is the average of the proportion of dirty backward citations among the total backward citations (Sectoral average of intensity metric). Overall, the figure shows that the range and intensity of direct connection tend to be positively correlated, confirming that our D metric well represents the connectivity between clean and dirty technologies.

Supplementary Figure 8. The pattern of sectoral connections from clean patents to prior dirty patents



Notes. X-axis presents the mean of distances from patents in a certain clean sector (indicated by color of dots) to patents in a certain prior dirty sector (indicated by shape of dots). Y-axis presents the share of connected patents in a certain clean sector to patents in a certain prior dirty sector. (a) For all clean sectors. (b) For clean energy sub sectors only.