UNIVERSITY OF CAMBRIDGE

Faculty of Economics

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 2013

IDENTIFYING INNOVATIVE ACTORS IN THE ELECTRICITY SUPPLY INDUSTRY USING MACHINE LEARNING: AN APPLICATION TO UK PATENT DATA

G.G. Dolphin M.G. Pollitt

3 March 2020

The recent history of the Electricity Supply Industry (ESI) of major western economies was marked by two fundamental changes: a transition toward liberalised electricity markets and a policy-led push to decarbonise the electricity generation portfolio. These changes not only affected the pace and nature of innovation activity in the sector but also altered the set of innovative actors. The present paper provides a methodology to identify these actors, which we apply to priority patents filed at the UK Intellectual Property Office over the period 1955-2016. The analysis also indicates that (i) the recent increase in innovation activity originates overwhelmingly from upstream Original Equipment Manufacturers and (ii) innovation activity in `green' electricity supply technologies slowed down in recent years.

UNIVERSITY OF CAMBRIDGE | Energy Policy Research Group

Identifying innovative actors in the Electricity Supply Industry using machine learning: an application to UK patent data

EPRG Working Paper 2004

Cambridge Working Paper in Economics 2013

G.G. Dolphin and M.G. Pollitt

Abstract The recent history of the Electricity Supply Industry (ESI) of major western economies was marked by two fundamental changes: a transition toward liberalised electricity markets and a policy-led push to decarbonise the electricity generation portfolio. These changes not only affected the pace and nature of innovation activity in the sector but also altered the set of innovative actors. The present paper provides a methodology to identify these actors, which we apply to priority patents filed at the UK Intellectual Property Office over the period 1955-2016. The analysis also indicates that (i) the recent increase in innovation activity originates overwhelmingly from upstream Original Equipment Manufacturers and (ii) innovation activity in `green' electricity supply technologies slowed down in recent years.

Keywords innovation, electricity sector, machine learning

JEL Classification L94, O31, O38

Contact gd396@cam.ac.uk
Publication March 2020
Financial Support UK Economic and Social Research Council

Identifying innovative actors in the Electricity Supply Industry using machine learning: an application to UK patent data

G.G. Dolphin*[a] and M.G. Pollitt[a]

[a] *Cambridge Judge Business School and Energy Policy Research Group, University of Cambridge, Trumpington St, Cambridge CB2 1AG, United Kingdom*

This version: March 3, 2020

Abstract

The recent history of the Electricity Supply Industry (ESI) of major western economies was marked by two fundamental changes: a transition toward liberalised electricity markets and a policy-led push to decarbonise the electricity generation portfolio. These changes not only affected the pace and nature of innovation activity in the sector but also altered the set of innovative actors. The present paper provides a methodology to identify these actors, which we apply to priority patents filed at the UK Intellectual Property Office over the period 1955-2016. The analysis also indicates that (i) the recent increase in innovation activity originates overwhelmingly from upstream Original Equipment Manufacturers and (ii) innovation activity in 'green' electricity supply technologies slowed down in recent years.

Keywords: innovation, electricity sector, machine learning

JEL Classification: L94, O31, O38

*Corresponding author (gd396@cam.ac.uk). The author gratefully acknowledges the support of the UK Economic and Social Research Council.

1 Introduction

Over the course of the last thirty years, the Electricity Supply Industry (ESI) of several OECD economies experienced significant structural changes (International Energy Agency, 1999; Nicolli and Vona, 2019). These changes occurred against the backdrop – and under the impulse – of two pivotal policy developments. First, the liberalisation of their respective electricity sectors, which initiated a transition from vertical integration to unbundling of electricity supply activities (generation, transmission and distribution) and the introduction of wholesale competition.[1] Second, the development of increasingly stringent power sector decarbonisation policies, which at times came to co-exist with liberalisation agendas.

Liberalisation – and ensuing structural reorganisation of the ESI – had a significant impact on the innovation activity of its actors for several reasons. First, liberalisation of downstream stages of the ESI affected incentives to innovate of both downstream actors and upstream Original Equipment Manufacturers (OEMs).[2] Second, liberalisation changed the identity of downstream actors, from government-owned vertically integrated entities (and associated research centres) to private competitive firms, and hence altered the nature of the incentives their innovation activity is sensitive to. Third, these changes were often accompanied by a restructuring of public energy R&D institutions and a reduction in associated spending. In the UK, public spending on energy R&D (all technologies) decreased consistently between 1985 and 1999, only recovering from 2003 onward with increased funding directed at renewable electricity generation technologies. Finally, these regulatory changes occurred during – and allowed for – a period of increased internationalisation of ownership at every stage of the electricity supply chain, which affected the location of R&D activities. This was because while OEMs were already internationalised, firms ins the downstream stages of the electricity supply chain were initially domestic. When these firms also internationalised this induced a relocation

[1] This process was first initiated in the UK (Electricity Act 1989) and the US (1992 Energy Policy Act), and subsequently in the European Union (Directives 96/92/EC, 2003/54/EC, 2009/72/EC).

[2] Evidence based on US data suggest that downstream liberalisation led to a decrease in patenting activity of upstream actors (Sanyal and Ghosh, 2013).

of R+D activities (Jamasb and Pollitt, 2011).

With regard to changes in environmental policy stringency, jurisdictions around the world strengthened direct support for the development and deployment of renewable electricity generation technologies while at the same time increasing the (implicit or explicit) price on greenhouse gas emissions. In the US, while there was little policy initiative on the part of the federal government, several State legislatures introduced Renewable Portfolio Standards (see North Carolina Clean Energy Technology Center (2019), for detailed State-level information). In the EU, and hence in the UK, the first such policy was the directive on the promotion of electricity produced from renewable energy sources in the internal electricity market (Directive 2001/77/EC). The transformation of the UK generation portfolio is now well under way, with the share of electricity produced from renewable sources having risen from 3.5% in 2000 to 24.6% in 2016 (Eurostat, 2018) and 35.8% in the first quarter of 2019 (BEIS, 2019), allowing it to reduce the CO_2-intensity of the said portfolio from 480g CO_2/kWh in 2000 to 246g CO_2/kWh in 2017 (IEA, 2018).[3]

This paper sheds a descriptive light on the evolution of the characteristics of UK-based innovative actors in the ESI in the face of these structural changes, based on patent filings at the UK IPO over the period 1955-2016.[4] The focus on UK actors is motivated by both historical institutional developments and methodological constraints. Regarding the former, the UK has been at the forefront of key technological and policy developments, making it a particularly salient case-study for the purpose of our research. With respect to the latter, the scope of our study is limited by two factors. First, our patent selection approach involves the use of natural language processing techniques on patents' title and abstract and therefore requires to work in a single language (English in this case). Since patents have to be written in one of the official languages of the patent office at which they're filed, which may or may not include English, an extension

[3]Although the UK was slightly below the 2016 EU average (29.6%), this transition represents the 6^{th} largest increase among all EU member states over that period (Eurostat, 2018).

[4]At this stage, it should be noted that not all innovations are patented nor patentable (Jamasb and Pollitt, 2011) and hence patent filing counts should not be interpreted as providing an exhaustive account of innovation taking place with regard to specific technologies. However, to the extent that patent filings follow the trends in innovation activity, they provide an accurate proxy to capture them (Dechezleprêtre et al., 2011).

of this approach to patents filed at other patent offices is non-trivial. In addition, the linking with business structure database, which we perform in section 4.3, introduces an additional hurdle to cross-jurisdiction analysis in the sense that the universe of firms that such databases cover varies across jurisdictions.

We make the following contributions. First, we provide a patent search methodology that uses a supervised learning classification algorithm (random forest) to identify patents pertaining to electricity supply technologies. The classification is based on n-grams derived from the patents' title and abstract.[5] This approach allows us to address a standard shortcoming of keywords-based search, i.e. that the list of keywords is a subjective construction which might only partially account for the semantic field used by applicants to describe relevant inventions. In addition, it is flexible enough to allow identification of "lateral" innovation. Second, in contrast to a number of earlier studies – see section 2 – which focus on the impact of liberalisation and decarbonisation policies on aggregate innovation trends, we provide an in-depth discussion of the characteristics – and heterogeneity[6] – of (UK-based) actors carrying out innovation along the entire electricity supply chain, from OEMs to distribution companies. Third, compared to previous studies which tend to concentrate on generation technologies, it provides an industry-oriented perspective and broadens the technological focus so as to include all electricity supply technologies.

The approach taken in this paper provided us with important insights. First, the innovation activity shifted away from large (integrated) generation, transmission and distribution utilities to (smaller) equipment manufacturers or R&D firms. Patent filings by universities, although increasing as a share of total patent filings over time, remain marginal. Second, the distribution of patent filings over the sample period is heavily skewed, with a small number of actors constituting a large proportion of filings. This is particularly true for OEMs. Third, on a related note, we uncovered the predominant role

[5] An n-gram is a sub-sequence of n-elements constructed from a given sequence. In the case at hand in this paper, the sequence is a single string of (stems of) words (comprised of a patent's title and abstract) and the elements are the (stems of) words. Hence, for instance, a bi-gram is a string comprised of two (stems of) words.

[6] To our knowledge, only Noailly and Smeets (2015) discuss the role of firm heterogeneity in the context of electricity supply technologies.

played by lateral innovation in the development of fossil fuel electricity generation technologies (FF). Fourth, innovation in these technologies still represents a large proportion of yearly filings. Finally, with regard to UK-based OEMs specifically, the paper highlights a number of firm-level (technological) dynamics: (a) a majority of patents are filed by firms that are active in both fossil fuel and renewable electricity generation technologies (REN), (b) but 'mixed' firms have filed significantly more FF patents than REN patents; and only during the period 2007-2013 have these firms filed more REN than FF patents, (c) the increase in REN patent filings observed between 2005 and 2011 went hand in hand with an increase in the number of (small) technological entrants (i.e. firms patenting for the first time).

The rest of the paper is organised as follows. Section 2 reviews the relevant literature and introduces the analytical framework. Section 3 presents the construction of the dataset and section 4 analyses the dataset by (main) actors and technologies. Section 5 discusses policy implications. Section 6 concludes.

2 Innovation in the ESI – actors, technologies and incentives

Studying innovation activity at the sector level based on patent data presents a number of challenges. First, it calls for the identification of the intersection of *relevant* actors and technologies in order to construct the *relevant* set of patents – section 2.1. Second, given the diversity of innovative actors in the ESI, rationalising the observed patenting trends then requires a distinct discussion for each of them – section 2.2.

2.1 ESI actors and technologies

From an institutional perspective, innovative actors in the ESI – just like in any other sector – can be seen as belonging to one of the following categories: private corporations, government-owned non profit entities (e.g. vertically integrated utilities such as those existing prior to liberalisation), universities and research centres, individuals. This classification, introduced by the Centre for Research & Development Monitoring (2017), is used in section 4.1.

From an industry perspective, the ESI is usually understood as comprising an upstream stage (OEMs) and downstream stage (generation, transmission and distribution operators). In its investigation of innovation activity in the sector, prior literature followed this dichotomy and studied innovation by upstream equipment manufacturers and downstream generation, transmission and distribution entities separately. This is in part a reflection of the difference in the nature of incentives to innovate faced by actors in each stage of the ESI. The overwhelming majority of earlier studies focus on generation technologies and, as a result, mostly discuss innovation by upstream equipment manufacturers. Relatively fewer studies have investigated innovation by downstream actors; notable exceptions are Jamasb and Pollitt (2011, 2015) in the UK context.

One difference between upstream and downstream actors is that the latter are likely to have a narrower technological focus (i.e. on electricity supply technologies) whereas equipment manufacturers may have a more diversified innovation portfolio. Hence, unless these actors focus solely on electricity supply technologies, one cannot consider all patent filings by OEMs as pertaining to these technologies. Identifying filings that are specific to them requires us to filter by specific keywords or (IPC/CPC) technological codes.

This is why prior literature examining the innovation activity in the ESI has mainly worked based on the identification of key technologies and associated technological codes.[7] In addition, these studies focused primarily on electricity generation technologies. Johnstone et al. (2010) identifies IPC codes pertaining to renewable electricity generation technologies whereas Lanzi (2010) develops a methodology whereby IPC codes for both general and efficiency-enhancing fossil fuel generation technologies are uncovered. Taken together, they provide a comprehensive list of IPC codes pertaining to electricity generation technologies. The present study contributes to a more complete identification of electricity supply technologies by singling out IPC codes related to transmission and distribution technologies as well as other ESI-relevant technologies.

[7] Although the construction of patent datasets based on such an approach would in theory allow for a discussion around the actors by which they are filed, most studies do not include such a discussion.

2.2 Innovative actors and innovation incentives

Over the period under study in this paper, innovation in electricity supply technologies was primarily carried out by three distinct groups of actors: public R&D institutions, integrated utilities and private firms. Each faced different constraints and incentives and played a different role in the development of electricity supply technologies, which we briefly review in this section.

2.2.1 Public R&D institutions and integrated utilities' energy research

The development of electricity supply technologies over the second half of the XXst century owes much to the innovation activities carried out by integrated utilities and public R&D institutions. Indeed, some electricity generation technologies (e.g. nuclear) were developed through dedicated institutions, which arose from a commitment by public authorities to develop them. Such was the case, for instance, of the UK Atomic Energy Authority (UK AEA). The Authority, which initially oversaw the entire UK nuclear program, retained responsibility for solely research activities after a restructuring in 1971 (Atomic Energy Authority Act 1971). As highlighted by earlier literature (Jamasb and Pollitt, 2008, 2011) and as evidenced by our patent filing sample – see section 4.1 – the UK AEA played a prominent role in developing civil nuclear energy technologies as well as other related technologies. Furthermore, Jamasb et al. (2008) noted that the decision to break-up the energy laboratories previously operated by the UK AEA disregarded energy research policy considerations and was mostly a "side-effect" of competition policy. This had unfortunate consequences for the UK energy research activity, both public and private, since such large research bodies were also triggering innovation by private (smaller) entities.

2.2.2 Firms

Theoretical and empirical research into the drivers of innovation at the firm level suggest that (a) the competitive environment affects innovation incentives (Arrow, 1962; Gilbert and Newbery, 1982), but the relationship is non-monotonic (Aghion et al., 2005);[8] (b) relative input prices and the policy environment can affect the direction and pace of innovation (Hicks, 1932; Acemoglu, 2002); (c) firms' innovation patterns (i.e. intensity and quality) are heterogenous and depend on structural industry or firm-level factors (Schumpeter, 1942; Mansfield, 1962; Kamien and Schwartz, 1975).

Several studies investigated these propositions in the context of the (UK) ESI. Regarding the liberalisation process, most of them investigated the impact of such reform on the actors directly affected, i.e. generation, transmission and distribution operators. In the UK, Jamasb and Pollitt (2008, 2011, 2015) note a substantial decline in R&D in the electricity sector following liberalisation, which they attribute mainly to: (i) the positive correlation between public and private R&D spending in the UK electricity sector and the fall in public R&D over the liberalisation period; (ii) the fact that intensity of innovation activity is related to the (expected) payoff of innovation (Nemet, 2009) – by inducing competition among actors with low(er) market share, it reduced the market share of each individual electricity generator, thereby reducing the incentive to innovate.[9]

As Sanyal and Ghosh (2013) showed, the introduction of competition in the downstream generation sector also affected the innovation activity of (upstream) equipment manufacturers. They show that following the Energy Policy Act, patent applications by OEMs at the US PTO declined substantially. Building on the theoretical framework provided by Aghion et al. (2005), they propose that this net decline is the result of "a

[8] Schumpeter (1942) initially argued that (near-)monopoly firms in highly concentrated industries would have higher incentives and be better able to provide innovation than small competitive firms whereas Arrow (1962) pointed out that, owing to several market failures, the private provision of knowledge would fall short of the (socially) efficient level, regardless of the market structure. A later investigation of the competition-innovation relationship suggested that this issue would be partly alleviated if the monopoly faced credible entry pressures (Gilbert and Newbery, 1982). Moreover, Aghion et al. (2005) highlighted that competition could have a different effect on innovation depending on the composition of the industry – if it is mostly populated by *neck-and-neck* firms, then increased competition will induce more innovation whereas if it is mostly comprised of *leaders-followers* then increased competition might reduce the incentives for followers to innovate and reduce overall innovation.

[9] However, the development of new abatement technologies is linked to the existence of a demand for such technologies. The introduction of climate policies supporting such a demand might have counteracted the negative effect of liberalisation. In that respect, Fischer (2008) develops a theoretical model showing that government support for emissions control R&D is only effective if there is at least moderate environmental policy in place to encourage adoption of the resulting technologies, suggesting that it is the combination of environmental and technology policies that leads most effectively to a technological transition.

negative pure competition effect outweighing the positive escape competition effect arising out of competition among the upstream [Electric Equipment Manufacturers], and the positive appropriation effect arising out of IPP entry downstream" (Jamasb and Pollitt (2011), p. 314). The existence of a relationship between the structure of the downstream generation market and innovation by upstream equipment manufacturers is to be expected given that power suppliers (i.e. utilities) purchase innovation from upstream equipment manufacturers.

Besides the market structure, earlier literature also showed that market incentives and environmental policies affect the direction and pace of technical change. Popp (2002) finds evidence that higher energy prices induce innovation in "clean" and energy-saving technologies whereas Porter and van der Linde (1995) were the first to suggest that environmental regulation can stimulate the firms' green innovation activity, with Jaffe and Palmer (1997) providing supporting empirical evidence to this claim.

The role of environmental and climate policies in shaping technological development was subsequently discussed in the more general framework provided by the literatures on endogenous growth and directed technical change, with the attention shifting towards the role of these policies in initiating and/or sustaining innovation in climate-friendly technologies (Jaffe et al., 2002; Newell et al., 2006; Popp, 2010). We learn from these literatures that innovation in the dirty or green product depends on the relative strength of the market size and price effects.[10]

Empirical evidence regarding some of the above mechanisms was provided by Aghion et al. (2016) and Calel and Dechezlepretre (2016), both using firm-level data. The former study highlights two interesting features: (i) that firms tend to innovate more in clean (and less in dirty) technologies when they face higher tax-inclusive fuel prices; (ii) that there is path dependence in the type of innovation (clean/dirty) both from aggregate spillovers and from the firm's own innovation history.

Finally, other advances in this strand of literature shed light on the heterogeneity of

[10]The former encourages innovation in the larger market, the latter encourages innovation for the market with the highest price. Since the market for the dirty good is currently the relatively larger one, there is a risk that the market size effect drives the economy towards innovation in the dirty sector.

innovators and as a result provided more precise indications about the nature of innovation dynamics. Klette and Kortum (2004) suggest that technological transitions can occur both through a shift of innovation activities within existing firms and through innovation entry and exit; Noailly and Smeets (2015) point out that the empirical literature in this line of research documents several key stylised facts about innovating firms. Among them are the observations that: (i) the distribution of R&D intensity among firms is highly skewed, (ii) large established firms are very active innovators but tend to focus on improving existing technologies, (iii) more radical innovations are the preserve of small and new entrants.

2.3 Analytical framework

The above literature provides valuable guidance for the analysis of innovation patterns by UK actors. First, innovation by private and public entities is correlated with levels of public R&D spending. Second, the set of actors performing innovation in the ESI and the relative weight of each type of actor is likely to have changed as a result of both liberalisation and environmental policy changes.

With the liberalisation of the ESI and the quasi-disappearance of public research institutions, most of the innovation activity – at all stages of the electricity supply chain – is carried out by private entities which, in turn, strengthens the need to understand the dynamics driving their innovation activity.

Therefore, after reviewing innovation by all actors identified in our sample we further characterise innovation by UK-based OEMs. By linking patent information with business structure data at the firm-level, we relate innovation activity to the firms' own knowledge stock, age and size. Furthermore, following Noailly and Smeets (2015), we provide a discussion of firms' technological heterogeneity, making a distinction between technologically mixed and specialised firms, i.e. firms specialising in a single (generation) technology or in multiple (generation) technologies. We expect mixed firms to be larger and older and the bulk of climate change mitigation technologies innovation to be provided by smaller, younger new entrants (since higher energy prices and environmental policy, which became more stringent more recently, should trigger 'technological entry').

3 Patent data selection: Identifying ESI-specific patenting activity

The discussion in the present paper is based on a sample of *priority* patent applications filed at the UK Intellectual Property Office over the period 1955-2016.[11] This choice is motivated by the focus of our study, which is (primarily) to identify UK-based innovation. Information related to these filings is extracted from the European Patent Office (2018) Worldwide Statistical Patent Database (version Spring 2018) via PATSTAT online.[12]

3.1 Patent search methodologies

Using patent filings for our purposes presents two challenges. The first pertains to the standard limitations of patent data (Calel and Dechezlepretre, 2016), which only capture part of the innovation activity (Jamasb and Pollitt, 2011), and hence require that the observed trends be discussed with due regard to the nature of the technologies at hand as well as the broader patenting context.[13] That is, we need to: (i) understand whether the trend observed at the industry level simply follows an aggregate trend in patenting or if there is indeed some industry-specific pattern; (ii) make sure that these filings continue to capture some of the firms' innovation activity, given that filing at the UK IPO is not the only route available to seek protection for a UK-based entity.[14] To see the former,

[11] This follows Jamasb and Pollitt (2011), except for the fact that their sample also includes patents whose priority country is the UK filed at the EPO and WIPO. However, these latter filings are bound to be 'duplicates' and hence provide little additional information with regard to our objective of identifying UK-based innovative actors.

[12] The SQL queries used to query the PATSTAT database are provided in appendix B.1.

[13] The first comprehensive account of the economic relevance and availability for research of patent data was given by Griliches (1990) but their use, which has grown dramatically over time as both the quality of patent statistics and their availability have increased, dates back to Bound et al. (1984). Furthermore, output measures of the innovation process are generally preferable to input measures such as R&D spending (Dechezleprêtre et al., 2011).

[14] Patenting at the UK IPO reached an all time high in 1969 (63614 filings) and decreased steadily until today (22072 filings in 2017). This includes direct and Patent Cooperation Treaty (PCT) applications. Yet, since the opening of the European Patent Office (EPO) in 1978, protection in the UK can be obtained via this route too. Total filings at the EPO were initially marginal (3598 in 1978) and became increasingly popular, especially since the early 1990s (60754 in 1990; 166585 in 2017).

Figure 1 shows both the absolute and relative (i.e. as a share of total UK IPO patent applications) count of patent filings. Over the studied period, both counts follow the same pattern, suggesting that the absolute count of patent filings at the UK IPO does indeed reflect some industry-specific pattern. As for the latter, there continues to be a "home-bias" which induces inventors to file the first (priority) patent to the intellectual property office that is closest to "home" (Dechezleprêtre et al., 2011). In addition, given that filing at a national office is cheaper than filing at a regional office, the former allows inventors to swiftly acquire a filing of which they can claim the priority when they file a patent at the latter.

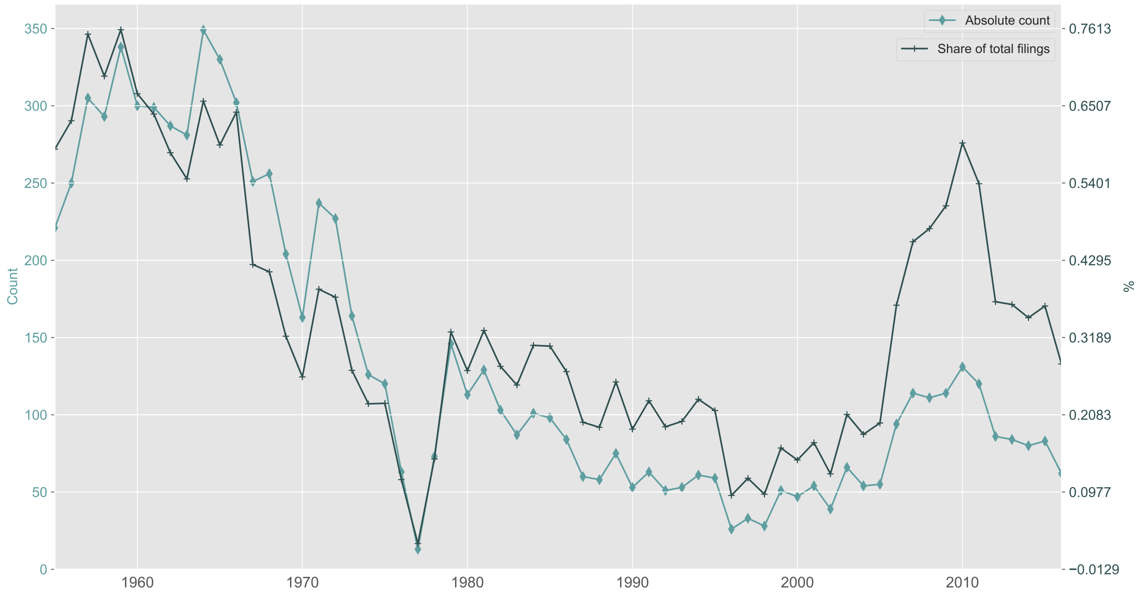

Figure 1: UK ESI patent applications, absolute count and share of total UK IPO filings

The second challenge arises because of a mismatch between the nature of our study – which investigates innovation trends at the sector level – and the structure of the patent classification system – which is based on technical features rather than sector of origin or "destination"(Hall et al., 2001; Jamasb and Pollitt, 2011). As a result, identifying the galaxy of patents "relevant" to a particular sector of the economy must continue to rely on *ad hoc* search strategies. These usually take one of three forms – actor-based, keyword-based and technology-based – or combinations thereof (Jamasb and Pollitt, 2011). In the first approach, the patent search is based on the name of relevant actors (e.g. utilities, equipment manufacturers, research institutes,. . .). This, however, must either rely on the researcher's prior knowledge of the actors' names or on international classifications nomenclature (such as ISIC – or its European equivalent, NACE) to identify the firms belonging to specific sectors – see, e.g., Bound et al. (1984). The former might leave out patents submitted by smaller (and less likely to be known) actors while the latter might leave out actors whose primary affiliation is not the sector under scrutiny. Furthermore, while the number of downstream actors is relatively limited and the UK Office for Gas and Electricity Markets (Ofgem) maintains a list of licensed generation, transmission and distribution companies, the number of equipment manufacturers is potentially much larger, which makes a search based on their names impractical.[15]

The second approach relies on a list of keywords (and combinations thereof) and can be used for sector and technology oriented patent search. This addresses some of the limitations of the actor-based search but introduces new ones (e.g. subjectivity in the choice of keywords, inability to cope with strategic 'naming' behaviour on the part of applicants).

Finally, a third approach consists in identifying the patents using their International Patent Classification (IPC) or Cooperative Patent Classification (CPC) technology codes. This approach has been adopted in Dechezleprêtre et al. (2011); Dechezleprêtre and Glachant (2014); Lanzi (2010) and has been the approach taken to establish the EPO-CPC climate change mitigation technologies classes.[16] It relies on identifying the codes associated with the technology(ies) under scrutiny (i.e. electricity supply), which in itself is not immune to errors and the accuracy of which is likely to be higher for well-established technologies than for nascent ones.

3.2 Our patent search strategy

We develop a patent selection strategy that addresses some of the limitations highlighted above, minimises the measurement error, i.e. inclusion of irrelevant patents and exclu-

[15] In the US, the Energy Information Administration maintains a list of equipment manufacturers – see Sanyal and Ghosh (2013) – but no such list exist for the UK and, even if it did, it might not be exhaustive.

[16] See https://www.epo.org/news-issues/issues/classification/classification.html for more information about this classification.

sion of relevant ones (Dechezleprêtre et al., 2011), and is suited to our objectives of (i) identifying UK-based innovative actors along the electricity supply chain, (ii) identifying their innovation activity, (iii) supplement IPC classes list with codes relevant to transmission and distribution, and other relevant technologies. Our strategy, which consists in a keywords-based search (KW) – combined with supervised machine learning (ML) classification – and an actors-based search, is described below.

We start with an initial dataset containing 346797 patent applications covering the innovation activity in the UK between 1955 and 2017.[17] This *core* patent dataset contains all patents with associated technology field(s) belonging to IPC categories 'B','F','G','H', with application authority GB and priority country GB over that period. Within that set, our patent selection starts with a keywords-based search on the patent title using keywords queries presented in Jamasb and Pollitt (2011) and aimed at covering electricity generation, transmission and distribution technologies. This search identifies 3072 distinct patent applications.

However, one drawback of such an approach is that it depends on a subjective keywords list, which may not be representative of the semantic field describing all the relevant technologies; it therefore may only partially capture the set of relevant patents. To address this concern, we resort to identifying relevant patents using a random forest classifier on a subset of our *core* patent dataset. This subset is the set of patents with IPC classification codes associated with the patents identified by our initial keywords-based search. In doing so, we hope to include patents that may be relevant to the ESI but that have been missed due to the use of words not included in our list to describe the patented invention. The rationale behind this approach being that the relevant patents that may not use the same keywords should still have been assigned the same IPC class. This step produces a sample of 59757 patent applications, which is bound to include some patents that do not pertain to electricity supply technologies. Therefore, the last step of our ML

[17] This is the total number of patents with an abstract AND a title. Our initial patent search, which was truncated to return only patent applications with an abstract, returned 354760 patents, 7963 of which did not have a title and were excluded from the sample. In addition, note that we downloaded filings up to 2017 but later truncated our sample to 2016 to account for a lag between filing and actual recording in the database.

approach is to distinguish between the patents relevant to electricity supply technologies and those that are not using our classification algorithm. The construction and training of the classifier is described in appendix A. It identified 3498 patent applications, 1811 of which had also been identified by the initial keywords-search. Given that scrutiny of the patents identified by the keywords search but not by the ML search indicate that most of them are relevant to electricity supply technologies, we keep the union of results of these two sets in our datasets.

This first set is complemented with the patents identified by an actors-based search, which targets only downstream ESI actors and relies on Ofgem's list of licensed electricity generators, transmission companies and distributors.[18] This list is complemented with entities identified in Jamasb and Pollitt (2011).[19] The search proceeds as follows. First, we search the PatStat Standardized Names in our original set for matches with entities in our list. Second, we perform a manual check and remove incorrectly identified patentees.[20] This leaves us with 24 actors and identifies 3731 distinct patent filings. The list of actors for which at least one filing was found is presented in appendix A.

Table 1 summarises the results of our patent search strategy, broken down by main category of applicants, and offers a comparison between the two search approaches. As it turns out, most of the patent filings identified by the ML-based search are by original equipment manufacturers (77%), followed by filings by individuals (15%), the Electricity Council (EC) and the UK Atomic Energy Authority (2%) – UK AEA, universities (0.5%) and integrated utilities. The actor-based search, on the other hand, focused on downstream ESI actors together with some actors known to have played a significant role in the technological development of the UK ESI (e.g. the Electricity Council and the Atomic Energy Authority).[21] The overwhelming majority of patents identified by this search

[18]This list is publicly available through the Electronic Public Register, accessible at https://epr.ofgem.gov.uk/Document and contains all documents related to licenses granted under the Gas Act 1986 and the Electricity Act 1989.

[19]Note that we also performed a search based on the names of the Global Ultimate Owners of the entities present in Ofgem's list but that few – if any – of these patents were associated with electricity supply technologies. Global Ultimate Owners were therefore excluded from our "actors"' list.

[20]For instance, patent applications filed by 'BP CHEMICALS' are removed from the dataset as they are not related to electricity supply technologies.

[21]The UK AEA was created in 1954 and was at the time responsible for the UK's civil and military nuclear programme, contributing very significantly to innovation in nuclear electricity generation

was filed by the UK AEA (85%), the remainder of the filings being distributed between generation companies (6%), integrated utilities (5%), and transmission and distribution companies (0.5%). Taken together, the ML search and the actors-based search provide us with a dataset containing 8389 patents, of which the largest proportion was filed by OEMs (44%), followed by the EC and the UK AEA (38%), individuals (8%), generation companies (3%), integrated utilities (2.5%). The table also highlights the complementarity of the ML and actors-based searches as there is few patents that are identified by both of them. This suggests that the patents filed by the actors identified in this paper may make use of a (slightly) different semantic field than that used in the keywords-based queries and that constructed by our random-forest classifier.[22] In addition, we note that there is a significant difference in the filing activity of companies (especially OEMs and generation operators) and individuals. The former do, on average, file 5.9 patents over the period covered whereas the latter filed only 1.2 patents on average; suggesting that companies have more systematic and organised innovation activities leading to sustained patent filings.

4 Whose – and what – innovation?

Equipped with the sample of 8389 patents presented above we review the patent filing activity in the UK ESI over the period 1955-2016.[23] After a brief review of the general trend in patent application filings over this period, we first shed light on the actors – or actor categories – from which they originate; with the view of identifying the evolution of patent filings across all industry actors – section 4.1. Next, we analyse the (technological) nature of these filings, shedding light on the technological transition that occurred – section 4.2. Finally, for a subset of actors, i.e. UK-based OEMs, we relate patent filings

technologies. Following the Atomic Energy Authority Act 1971, only research activities remained with the Authority. The Electricity Council, on the other hand, was set up in 1957 and tasked to oversee the electricity supply industry in England and Wales. It maintained research activities throughout its lifetime, especially in fossil-fuel based generation technologies.

[22]Arguably, it would be possible to design the ML search and train the classifier on a different sample so as to increase the overlap. This would make sense if the researcher was interested in relying on a single type of search; it was not the avenue pursued in this paper.

[23]Given the existence of a lag between the reporting of patent filings by national patent offices and their inclusion in the PATSTAT database, we exclude the most recent year in the sample, 2017.

Table 1: Patent searches summary (1955-2016)

	Actor type	ESI stage	KW	ML	(KW ∪ ML)	Actors	(KW ∪ ML) ∩ Actors	KW ∪ ML ∪ Actors
N. Patents	Companies	OEM	2364		3677	-	0	3677
		Generation	2		3	279	3	279
		Transmission	-		0	3	0	3
		Distribution	-		0	12	0	12
	Integrated utilities		10		11	222	11	222
	Universities		19		26	-	-	26
	Individuals		496		696	-	-	696
	EC & UK AEA		20		87	3189	87	3189
	Other		160		258	24	0	282
	All actors		3072	[3498]	4759	3730	101	8388
N. applicants/ assignees	Companies	OEM	456		658	-	-	658
		Generation	2		2	10	2	13
		Transmission	0		0	2	0	2
		Distribution	0		0	4	1	4
	Integrated utilities		1		1	3	0	4
	Universities		14		18	-	-	18
	Individuals		428		571	-	-	571
	EC & UK AEA		1		1	2	1	2
	Other		88		128	3	0	131
	All actors		990		1379	24	3	1400

The number of applicants in the table above is based on an author-created unique entity identifier. It differs from the number of distinct 'psn_id"s associated with the identified patents since, at times, several of them refer to a single legal entity. Some patents that have been manually removed (e.g. motor vehicle internal combustion engine)

to firm-level business structure data – section 4.3.

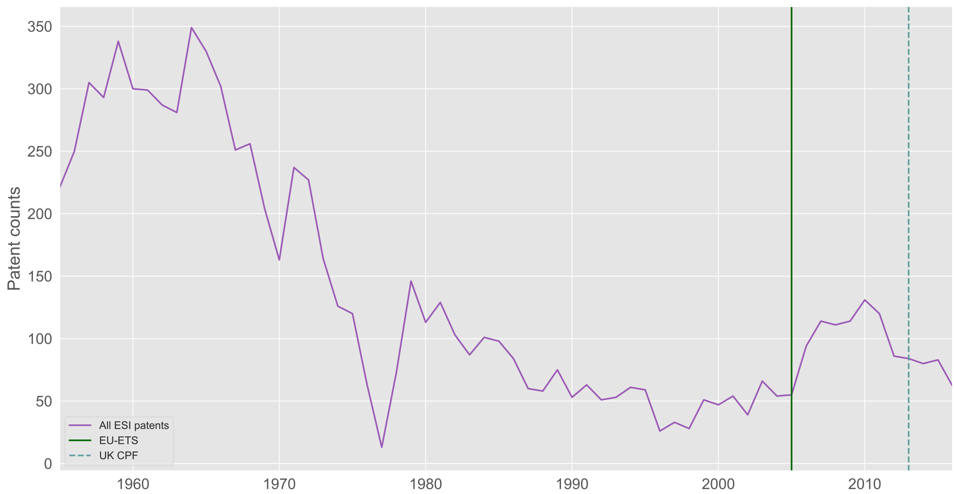

Figure 2: Patent applications at the UK Intellectual Property Office, 1955-2016

Aggregate trends are apparent in Figure 2, where we observe a clear decrease in total patent filing activity until the late 1990s, at which point an increase in filings relating to climate change mitigation technologies (as classified by the European Patent Office (2013)), brought about a revival in patenting. Put differently, yearly filings at the UK

IPO averaged 244 patents/year between 1955 and 1976, and just 74 patents/year in the subsequent period.

4.1 Origin of patent filings

The aggregate trends observed above can be broken down according to the actors from which the filings originate. First, we distinguish between the main categories of patent applicants, as identified in section 2.1. Figure 3 presents the yearly patent filings introduced by each type of applicant, as a share of total applications in our sample. It clearly highlights the importance of three categories of actors: 'Company', 'Government non-profit' and 'Individual'. These categories account for 47%, 41%, and 8% of filings in our sample, respectively. In addition, the implications of the liberalisation and dismantling of vertically integrated utilities is indicated clearly by the change in the relative importance of patents filed by 'companies' and those identified as 'government non-profit' organisations at the start of the 1990s.[24] It also makes apparent the rise in importance of patent filings by 'individual' applicants which have mostly filed patents pertaining to REN technologies.

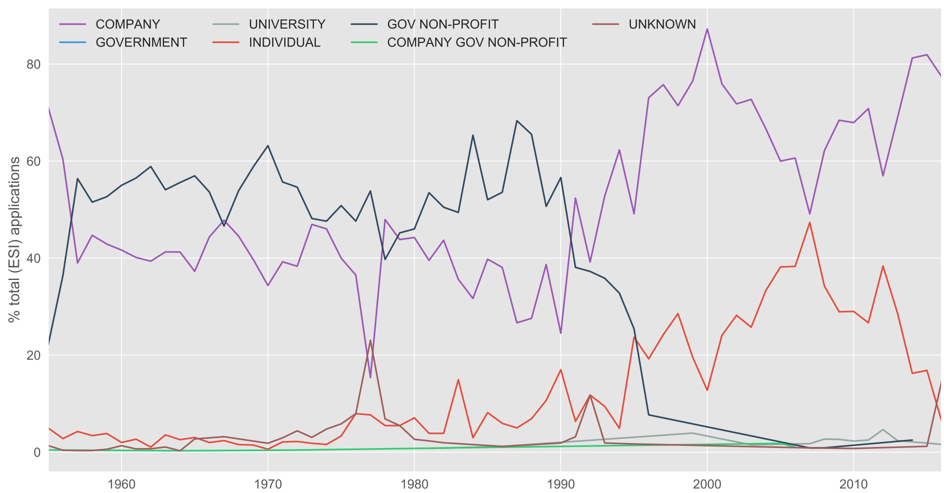

Figure 3: UK ESI patent applications, by type of applicant

[24]This relative change in the origin of patent is somewhat "mechanical" and likely reflects the transfer of assets previously owned by vertically integrated utilities to private corporations.

Our second categorisation distinguishes between actors along the electricity supply chain. As alluded to earlier, we identify upstream OEMs and downstream generation, transmission, and distribution companies as well as vertically integrated entities (e.g. Central Electricity Generating Board) and two key actors of the UK ESI, the Electricity Council and the Atomic Energy Authority – see Figure 4. A striking feature of the picture painted by this figure is the predominant role played by OEMs. They were responsible for a significant share of total yearly filings (on average, 105.6/year between 1955 and 1977, 20.7/year between 1978 and 2000 and 52.9 between 2001 and 2016), the rest of it originating primarily from the UK Atomic Energy Authority and the Electricity Council.

From 1978 onward, patent filings by OEMs decrease slightly faster than those of downstream actors, altering the relative importance of each type of actors' contribution to total patent filings. Patent applications by generation, transmission and distribution actors at the UK IPO remained strong until the late 1990s – which corresponds to the full roll out of the provisions of the UK Electricity Act, while innovation activity by equipment manufacturers started dwindling as soon as the early 1970s. Interestingly, the patenting activity of OEMs remained stable throughout the liberalisation period and started increasing again towards the late 1990s. By the mid-1990s, OEMs represented again about half of total patent filings and, as patent filings by downstream actors almost vanished from 2002 onward, it represented an ever larger share of filings, accounting for most of the recovery in patent filings. Overall, insights provided by Figure 4 suggest that (i) original equipment manufacturers have always played a significant role in patenting activity, (ii) the relative importance of this activity has grown in recent years as patent filings by downstream actors dwindled to extremely low levels.

This aggregate picture, however, hides a more subtle feature: the distribution of patent filings (among actors) is heavily skewed. This observation matches a well known trait of patent applications: they are concentrated within the hands of a few key actors, both at the country – most patents are filed in a small number of offices[25] – and sector level –

[25]The so-called IP5 group, comprised of the US Patent and Trademark Office (USPTO), the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), and the National Intellectual Property Administration (CNIPA formerly SIPO) in China.

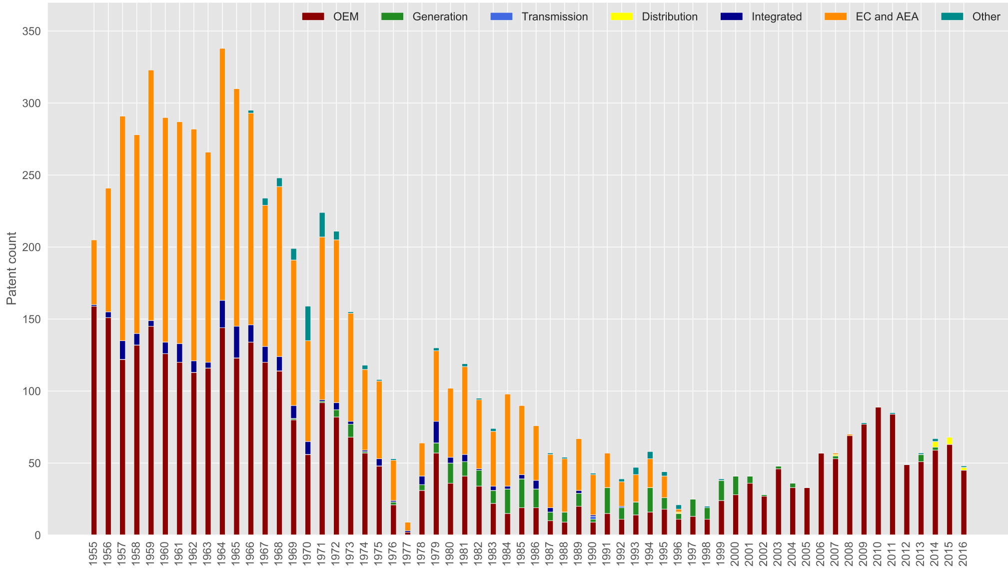

Figure 4: Annual patent filings in the UK ESI, by type of actor

within each sector, a few key players concentrate most R&D activity and patent filings. This is apparent in Figures 5 and 6, which present patent applicants over the period 1955-1990 (prior to liberalisation) and 1991-2017 (post-liberalisation), ranked in decreasing order of number of patent applications filed. We make a number of observations. First, quite unsurprisingly, the UK AEA tops the ranking over the period 1955-1990. Second, more interesting is the fact that Rolls-Royce has filed the second largest number of patents over the period 1955-1990 and the first largest over the period 1990-2016.[26] Over the entire period covered in our sample, it accounted for 41% of patents filed by OEMs and 18% of all filings identified over the period. This observation is particularly interesting given that a number of Rolls-Royce's patent filings pertain to jet engine turbines rather than turbines specifically destined to be used in electricity generating power plants – see next section for further discussion. These filings nonetheless do bear relevance to electricity generation technologies to the extent that, as noted by Joskow (1998) (p.50), pivotal "innovations in CCGT technologies [drew] on complementary research on the

[26] All of Rolls Royce's activities were part of a single entity until 1971, at which point its motor car activities were split from its aerospace, power systems and defence activities. The latter became part of a new entity, Rolls Royce plc. Figures 5 and 6 show the filings of the latter.

development of jet engines for commercial aircrafts". This also explains the presence of entities like Power Jets (R&D) and Bristol Siddeley Engines among entities with the largest number of patent filing in this sample.

Turning to the post-liberalisation period, we observe the effect of both the dismantling of integrated utilities and the emergence of their privatised successors as innovative actors – with patents filed by National Power, Drax Power – and, among OEMs, the emergence of actors focusing on renewable technologies, especially wind – with about a hundred patents filed by the Danish company Vestas.

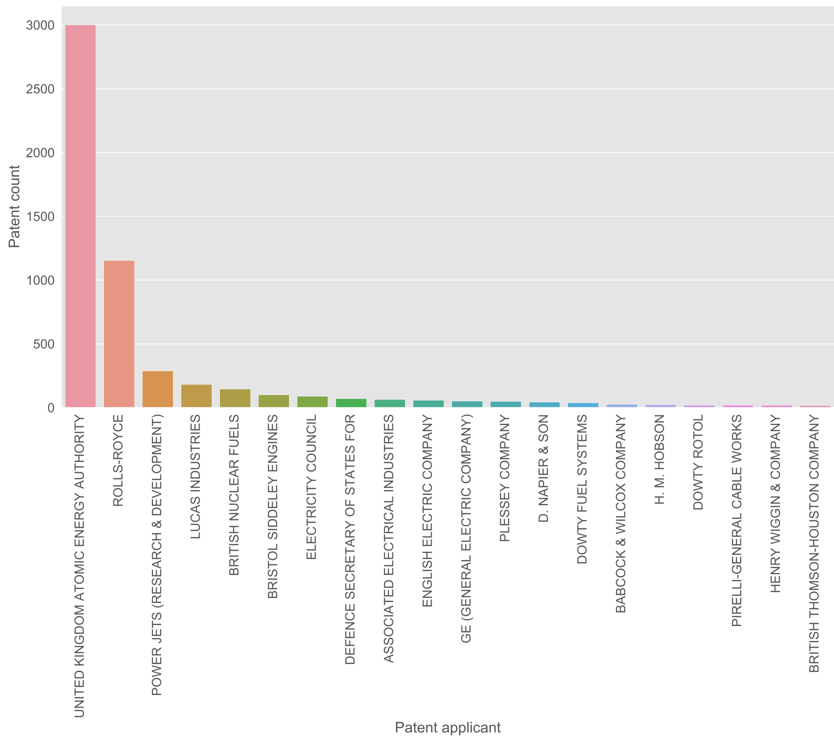

Figure 5: Patent filings, by assignee – 1955-1990

Finally, given the existence of a "home-bias", one would expect most patent filings in a dataset constructed based on priority filings at the UK IPO to have been made by UK-based applicants. While this is indeed the case, we note that some of the patent filings in our sample originated from non-UK actors, as observed in Figure 7.

This is mostly the case among OEMs, which have historically operated across national markets and sought protection for their innovation in their non-domestic/ export markets; whereas downstream actors remained focus on their domestic markets, especially until the liberalisation of the sector. For this former category of actors, in all years between

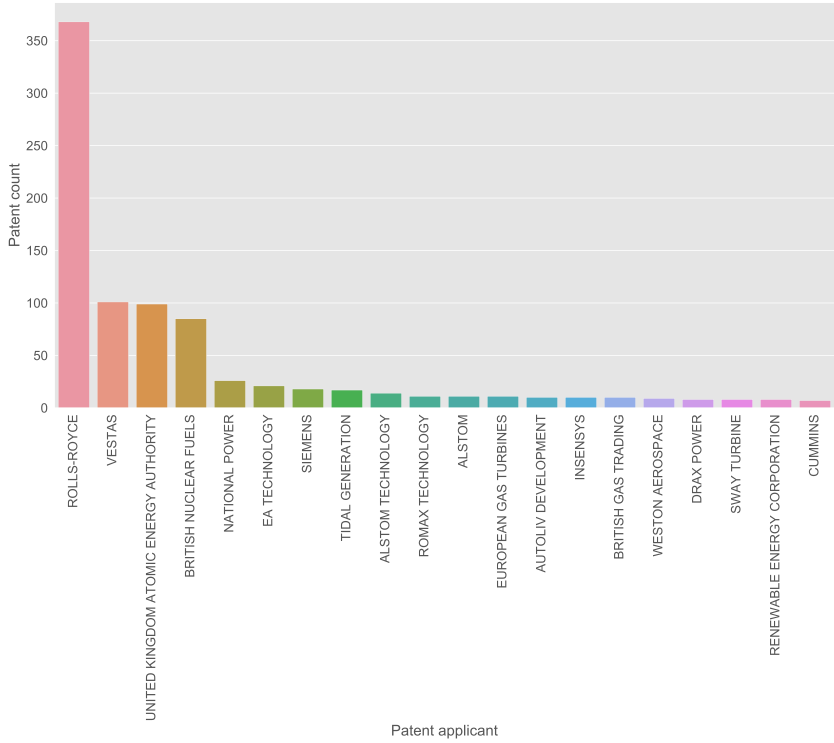

Figure 6: Patent filings, by assignee – 1991-2016

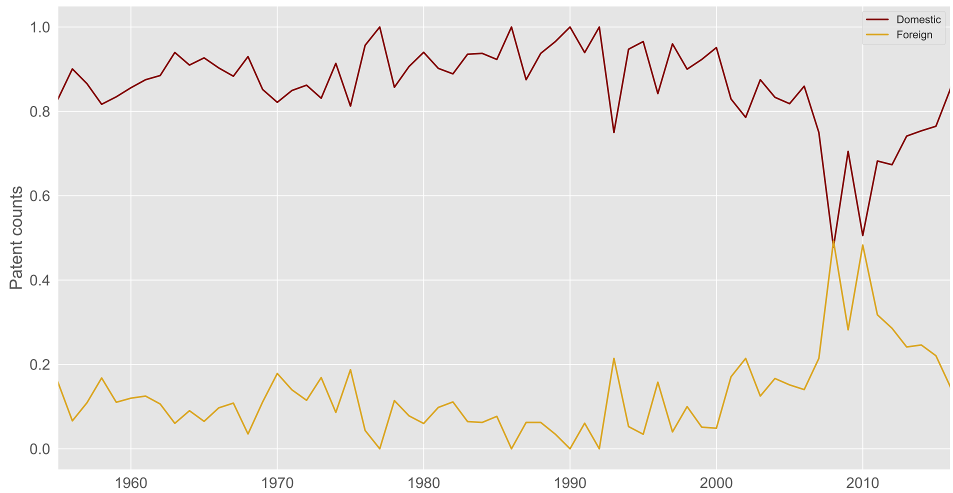

Figure 7: Share of patent filings by UK and foreign OEMs

1955 and 2000 (excepted 1993), the proportion of priority patents filed at the UK IPO by UK applicants was above 80%. This proportion declined steadily between 2000 and 2010, and recovered thereafter, which suggests that the increase in filings by OEMs observed between 2008 and 2011 was partly due to the activity of entities based outside the UK.

4.2 Nature of patent filings

The discussion in the previous section sheds light on the (main) actors which have filed patents over the period under study. We now look at the type of (electricity supply) technologies to which they pertain. We make two main distinctions. First, between generation, transmission and distribution technologies. Second, within generation technologies, between renewables (REN), fossil-fuel (FF) and efficiency-enhancing fossil-fuel technologies (FF-E), and nuclear (NUC).

In order to allocate patents to specific technological categories, we rely on (IPC and CPC) technological codes. Depending on the technology at hand, these codes are identified either based on earlier literature or on our own research. Earlier literature provides the IPC or CPC codes related to REN (Johnstone et al., 2010; European Patent Office, 2013), FF & FF-E (Lanzi, 2010), and NUC generation technologies (European Patent Office, 2013). On the contrary, technology codes pertaining to transmission and distribution technologies or other ESI-related technologies have been less documented. To identify these codes, we proceed as follows. First, we read and review some of the patents in our sample and assign them to specific technological categories (generation, transmission & distribution, energy storage, other) and sub-categories (e.g. type of generation technology, core technology vs. manufacturing processes). Second, we identify, for each technological group, all associated IPC/CPC 4-digit classes, ranked in descending order of attribution (i.e. the class with the highest number of occurrences is listed first). Finally, we check the first classes to determine whether or not they relate to the technology at hand. This leaves us with a set of technology codes pertaining to our technologies of interest. Table 7 in appendix C provides a complete list of IPC/CPC codes used to classify technologies in this paper.

This investigation confirms that not all patents in our sample are, strictly speaking, related to electricity supply technologies as identified by previous literature. In particular, the sample contains patents related to jet propulsion engines (and mounting thereof), instruments of measurement (e.g. radioactivity detection, utility metering,...), manufacturing processes of engines and turbines, general engineering and pollution control equipment. This is the case for two reasons. First, the actor-based search identifies patents by their applicant's name and therefore disregard their technological aspect. Second, the ML search was designed in such way that some *closely related* technologies would be identified.

In our sample, 2841 (37%) of patents relate to generation technologies and 349 (4.5%) to transmission and distribution. Within generation technologies, 428 (15%) patent filings pertained to renewables, 887 (31%) to fossil fuel generation technologies, 287 (10%) to efficiency enhancing fossil-fuel generation technologies, and 1239 (44%) to nuclear energy. In addition, we note that patent filings in jet engine technologies (1633) and instruments (1091) accounted for 22% and 14% of total filings in our sample, respectively.[27]

Figure 8 presents the evolution of such filings and confirms that the majority of filings was directed at generation technologies, with very few filings pertaining to transmission and distribution technologies, except in the periods 1955-1965 and 2005-2015. Somewhat surprisingly, filings for efficiency-enhancing fossil fuel technologies remained low throughout the period under study. One also notes that the decline in patent filings relating to nuclear power since the mid-1960s only partly explains the decrease in total filings, especially since patent filings for renewable technologies remained fairly stable over that period.

Note that the paucity of filings for innovations pertaining to transmission and distribution of technologies might not accurately reflect the innovation activity in those technologies as most of it has been incentivised through Ofgem's Electricity Network Innovation Competition, which includes a requirement that innovation outcomes be disseminated and made available to other parties (Ofgem, 2017).

In addition to the flow of filings presented above, we can also analyse the evolution of the industry *knowledge stock* over time, giving and indication of the knowledge base present in the UK ESI with regard to specific technologies. We focus on generation technologies. Figure 9 presents the discounted cumulative knowledge stock (proxied by

[27]For consistency with the data presented in figures 4 and 8, these figures exclude patent filings by individuals (696). Hence shares are calculated with regard to 8389-696 = 7693. Individuals have primarily filed patents in renewable generation technologies; including such patents would tilt the reported shares toward these technologies.

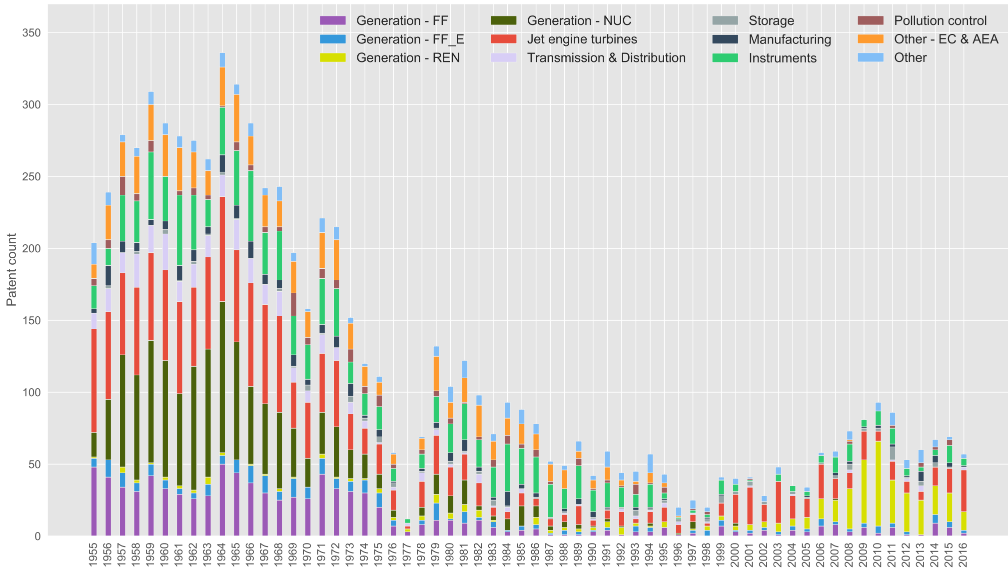

Figure 8: Annual patent filings in the UK ESI, by type of technology

the cumulative number of granted patents) of the UK ESI, using a 15% discount rate across all technologies (Hall and Mairesse, 1995).[28] We note that there is a steep increase in the industry's patent stock between 1960 and the early 1970s, primarily due to the increase in the stock of patents related to (i) fossil-fuel (ii) nuclear generation technologies. The stock of REN patents initially only rose very slowly, with the pace of increase rising slightly only toward the late 1970s. Interestingly, the value of the REN stock does not overtake that of NUC before the early 2000s (1990s if we include patenting by individuals) and not at all that of FF.

4.3 Actors' characteristics and innovative output

Identifying trends in patenting activity provides valuable insights into the direction and pace of technological change but falls short of shedding light on their micro-foundations and, in particular, the heterogeneity of actors driving these developments. Building on the discussion in section 4.1, we analyse further these developments by matching patent

[28]Given that our dataset starts in 1955 and that we don't hold any information about the stock of patent filings prior to that year, we truncate the time series and disregard the first five years of our sample, presenting the evolution of the stock from 1960 onward.

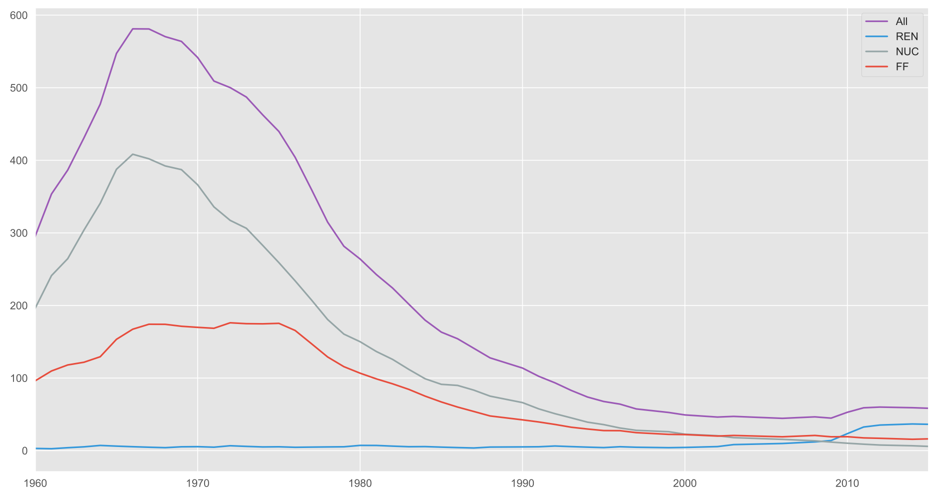

Figure 9: UK ESI discounted cumulative stock of granted patents, 1960-2016

filings with their corresponding legal entity using the patentee information associated with each application (provided in the PATSTAT database). The patent filing activity is then analysed in relation to filing history and firm business structure data such as age (using the date of incorporation) and size (number of employees and/or turnover). Information relating to the filing history is based on the patent sample identified above while business structure information is taken from the Bureau van Dijk (2018) FAME database.[29]

4.3.1 Matching

Matching patent data with financial data requires that the patents be associated with the correct legal/financial entity. As highlighted by previous literature, this matching is rendered difficult by the fact that the recorded patent applicants differ from business entities. Indeed, a given patenting entity may: (i) file applications under slightly different names (sometimes because of legal name change), (ii) apply under a name different to the corresponding legal entity, (iii) be a subsidiary (or plant) of a mother firm.

Regarding these issues, the OECD led an effort to (i) harmonise patent assignee names

[29]Note that the patent applications data cover the period 1955-2016 whereas the business structure data only cover the period 1997-2016. In addition, our database contains the date of incorporation of each legal entity.

(Magerman et al., 2006) and (ii) link patent assignees with business entities. The former resulted in the creation of harmonised names for patent assignees (HAN) – PatStat Standardised Name (PSN) and associated ID – while the latter led to the creation a commercially available database – ORBIS-IP, Bureau van Dijk– containing both accounting and patent data. The harmonisation of patent assignee names did not, however, remove all duplicate entries (in some cases, multiple PSN's continue to refer to a single legal entity). Moreover, the standardised names do not necessarily correspond to the latest legally recorded name of the corresponding legal entity. Hence, in the absence of a common identifier linking patenting and legal entities (Bound et al., 1984; Torrisi et al., 2010), matching patent and business structure databases remains, despite recent advances, a non-trivial problem.

The researcher is thus faced with the following choice regarding their overall matching strategy: adopt and automated matching procedure based on secondary identifying features such as company names and postcodes present in both databases or manually assign an identifying number to the patentees that is also present in the business structure database (e.g. company registry number).[30] In both cases, the aim is to match all identified patentees with (at most) one legal entity identifier.

Given that we do not have (bulk) access to the FAME database data, we resort to a version of the latter option. First, using table tls207_pers_appln of the PatStat database, we associate the patent filings in our dataset with their patenting assignee, corresponding standardised names and id number as well as postal addresse(s). Next, we associate (each of) them with their corresponding Company Registration Number (CRN), retrieved from the UK Companies House's website. The assignment makes use of information on entity name and postcode obtained from the Centre for Research & Development Monitoring (2017) Person Augmented Table. However, since the address information contained in the table did not record the latest address of some of the legal entities, it was necessary

[30]The former approach is possible when working with a database like FAME (or ORBIS), which contains current and past names of legal entities but is not appropriate for Business Structure Databases (business registers) of National Statistics Offices, which usually contain an anonymised identifier rather than an entity's name. See earlier literature, e.g. OST (2014) execute a fuzzy matching between the EPO patent standardised names and Bureau van Dijk firm-level datasets (European Patent Office, 2018; Bureau van Dijk, 2018) on a key combining both of the above features.

to update that information using Companies House's register information. This allowed to find correspondences between past and current business register addresses and, given that, a patentee's harmonised name and a company's registration number.[31]

Table 2 summarises the matching for applicants identified as 'companies', leaving out individual applicants and government non-profit organisations. From Table 1, we recall that there were 677 company applicants. Among those, 428 were identified as UK-based applicants, 180 as foreign applicants, and 69 remain unidentified. In terms of patent filings, this means that we were able to match 2925 (95%) of the patents filed by UK OEMs,[32] all of the patents filed by UK electricity generation (279) and transmission (3) companies, and all of the patents filed by UK distribution companies (12). Finally, note that 604 (16%) patents filed by OEMs were so by foreign entities, and 148 patents were filed by applicants that could not be identified either as a UK or a foreign company.

Table 2: UK patents/applicants matching summary

	ESI Category	Matched patents	Matched applicants
Count	OEM	2925	411
	Generation	279	13
	Transmission	3	2
	Distribution	12	3
	All actors	3241	428
Share*	OEM	0.8	0.62
	Generation	1	1
	Transmission	1	1
	Distribution	1	0.75
	All actors	0.88	0.63

* Share of total number of 'COMPANY' applications or applicants.

4.3.2 Innovation by UK OEMs & business structure

In section 4.1, we established that the relative importance of OEMs in filing activity had grown over time, and especially so since 2002. We therefore seek a further understanding

[31] As alluded to earlier, some of the standardised name entries identified at this stage refer to the same legal entity and are therefore associated with the same CRN. We aggregate at the firm (i.e. legal entity)-level, retaining the CRN and the entity's most recent name and each entry is then associated with its patent portfolio and business structure variables. The matching results in a mapping file which records a legal entity's psn_id, psn_name, Company Registration Number. It also records the type of ESI actor (OEM, Generation, Transmission, Distribution) and whether it was identified by the keywords- or actors-based search.

[32] This represents 80% of patents filed by all OEMs.

of the characteristics and innovation dynamics of these actors, focusing on those whose activities are located in the UK.[33] In particular, we investigate the patterns of technological entry and exit (Malerba and Orsenigo, 1999), the "technological inertia" (path dependence) that characterises patent filings at the firm-level, and their relationship with two key firm structure characteristics, age and size.

Following Noailly and Smeets (2015), we distinguish between technologically mixed firms – which innovate in at least two types of electricity generation technologies – and specific firms – which innovate in only one of them. This latter classification is based on the composition of the cumulative patent portfolio of the firm in the last year of the sample (2016). Technologically heterogeneous firms are labelled 'mixed' whereas technologically specialised firms are labelled 'green', 'brown' or 'nuclear', depending on whether their cumulative patent portfolio contains only REN, FF or FF_E, NUC patents, respectively. Firms that do not file patents in generation technologies but do patent in other technologies are labelled as 'other'. Within our set of UK OEMs, 36% of the firms that have patented are 'green' firms, 12% are 'brown' firms, 4% are 'nuclear' firms, 4% are mixed firms. The remaining firms (46%) filed patents only in non-generation technologies.

Figure 10 shows the number of patent filings by each type of firm. Filings by technologically mixed firms have consistently outstripped filings by either their brown or green counterparts (except in 2007-2010 and 2012-2013). Given that these constitute only 3% of the firms in our sample, it suggests that, on average, they have a larger patent portfolio than technologically specialised ones. Moreover, Figure 11 suggests that this portfolio is skewed towards FF and FF-E electricity generation technologies: REN filings by mixed firms remained below the number of filings for FF and FF-E generation technologies in every single year in the sample and are extremely few. Interestingly, these filings exhibit an extremely strong correlation with patent filings pertaining to jet engine turbines, which further supports our claim that the technological development of FF technologies was 'complementary' to an existing knowledge base in the UK industry.

[33] We leave out downstream actors as their innovation activity has been the focus of prior studies, e.g. Jamasb and Pollitt (2011, 2015).

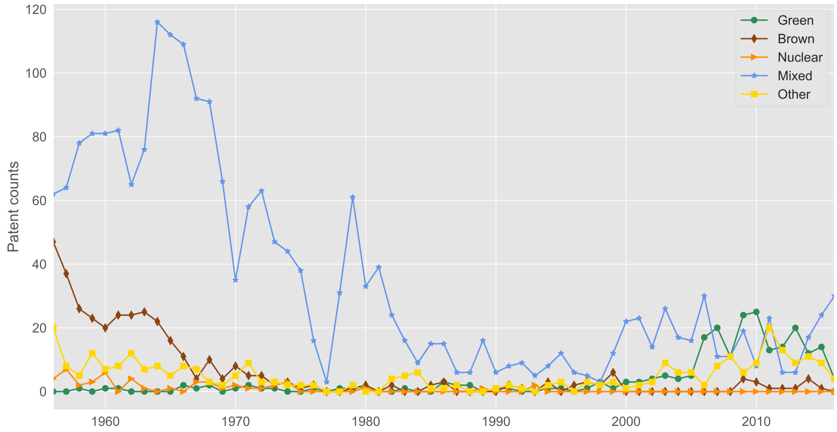

Figure 10: Patent applications by UK OEMs, by firm type

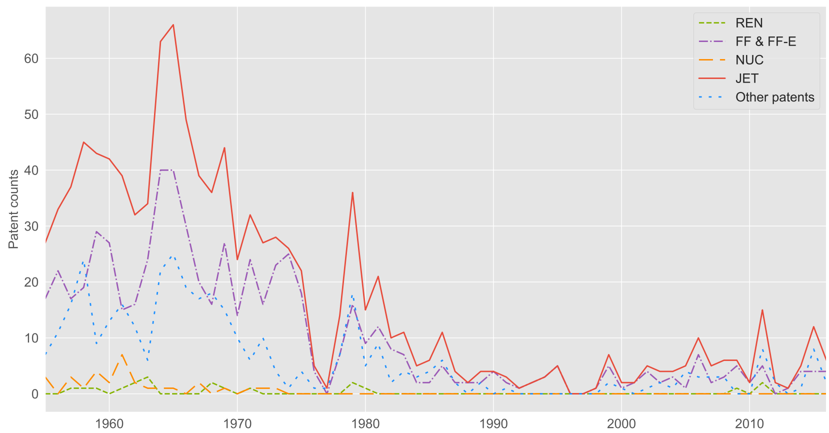

Figure 11: Patent applications by mixed UK OEMs

Technological entry

In light of the total number of REN patents filed by OEMs over the period, and especially in the years 2000-2010, the above observations suggest that the notable increase in patent filings for these technologies has been driven mostly by new (technological) entrants focusing specifically on them rather than by (older) mixed firms. This warrants a closer look at entry and exit dynamics.

The technological entry of a firm is defined as the first year in which it files a patent in the technological categories under consideration in our sample, regardless of its patenting history with respect to other technologies. Exit, on the other hand, is defined with respect to its discounted cumulative stock of patents: a firm is considered to exit the technological innovation market if the said stock reaches 0. It is calculated in the same way as at the industry-level using to the perpetual inventory method with a discount rate of 15% (Hall and Mairesse, 1995). Figure 12 presents the evolution of technologically active OEMs over the period 1955-2016.

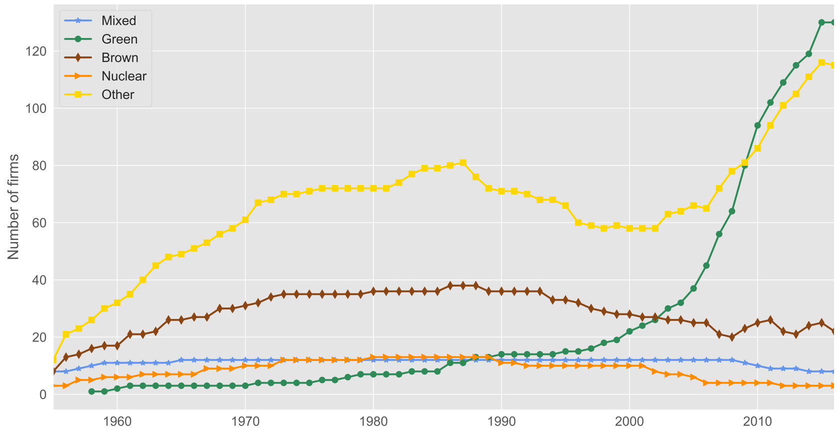

Figure 12: Technologically active firms, by firm type

The trends depicted indicate a rapidly increasing number of firms active in FF technologies until the early 1970s, corresponding to the development of fossil fuel fired power plants post WWII and public R&D funding for fossil fuel technologies – see figure 13. This increase is sustained – albeit more moderately – through the late 1980s, at which point the number of active brown firms starts decreasing steadily until it stabilises just above 20 in the late 2000s. The number of active mixed and nuclear firms follow a similar upward trend until 1980, at which point they both stabilise at 16 active firms. In the early 2000s, the number of active nuclear firms starts decreasing steadily. Finally, the number of active green firms grew steadily but slowly between 1958 and the mid-1990s, before experiencing an almost exponential increase from 1995 onward.

This is reflected in (technological) entry rates, which were significantly higher than the average during the years 2005-2011 for green firms than during the rest of the period. The average entry rate over the period 1960-2016 is 4.4% whereas it was 11.1% between 2005 and 2011. Given the large number of green OEMs filing patent applications since

2002 – above 5 in every single year – this implies that several of these applicants were new entrants, each filing on average a small number of patents. Table 3 provides further precision with regard to that observation. In every single year of our sample, green firms have indeed, on average, filed less patents than brown firms. In addition, there also seems to be a difference between firm type as mixed firms filed, on average, a higher number of both REN and FF patents than technologically specialised firms.

Table 3: Firm patenting activity: summary

Variable	Firm type	Mean	Median
REN patents	Green	0.023	0
	Brown	-	-
	Nuclear	-	-
	Mixed	0.025	0
	Other	-	-
FF patents	Green	-	-
	Brown	0.04	0
	Nuclear	-	-
	Mixed	0.67	0
	Other	-	-
Year of first REN innovation	Green	2005	2009
	Brown	-	-
	Nuclear	-	-
	Mixed	1972	1961
	Other		
Year of first FF innovation	Green	-	-
	Brown	1977	1968
	Nuclear	-	-
	Mixed	1963	1956
	Other	-	-
Year of first innovation	Green	2005	2009
	Brown	1974	1968
	Nuclear	1967	1972
	Mixed	1966	1961
	Other	1987	1994

Technological inertia, firm age and size

Finally, we relate these observations to some of the firms' own characteristics. First, the literature on directed technical change reviewed above suggests the existence of a path dependency in innovative activity – see also (Crespi and Scellato, 2015). Using firm-specific (discounted) knowledge stocks based on their patents filing history since 1955 (calculated, as in section 4.2, with the perpetual inventory method and a 15% discount rate), we investigate the correlation between a firm's knowledge stock and it's patenting activity. This correlation is positive across all firm types, and is highest for technologically mixed firms (0.84), followed by that for brown (0.66) and green (0.57) firms. The lower correlation observed for green firms is somewhat unsurprising given that most of them are recent innovators and have not been found to be the source of sustained innovation thus far.

Second, we investigate the relationship between innovation and firm age and size in our sample. While age (based on the date of incorporation) and innovation history are available for all years in our sample, our proxies for firm size (employees/turnover) are only available for the period 1997-2016. Even if a discussion of causal links between these variables and the probability of patenting in one or the other technological category is beyond the scope of this study, we can nonetheless provide some empirical evidence. In particular, we look at the value of these variables at the time of technological entry, broken down by firm type.

The evidence provided in Table 4 suggest that green and mixed firms were on average 9 years old at the time of their first patent filing whereas brown firms were significantly older, 14 years old on average. However, on average, green firms were smaller than brown firms but larger than mixed firms at the time of their first patent filing.

Table 4: Firm characteristics at time of first patent filing, by firm type

Variable	Firm type	Count	Mean	Std. dev.	Median	Min.	Max.
Firm age	Green	106	9.72	13.62	4.0	0.0	92.0
	Brown	14	18.57	18.5	8.0	0.0	50.0
	Mixed	3	34.0	3.61	33.0	31.0	38.0
	Nuclear	1	40.0	-	40.0	40.0	40.0
	Other	78	15.923	20.82	8.0	0.0	117.0
Employees	Green	20	715.25	2214.48	64.5	3.0	9989.0
	Brown	7	13156.29	29535	1633.00	55.0	80000.0
	Mixed	3	37633.33	2136.2	38500.0	35200.0	39200.0
	Nuclear	1	2614	-	2614	2614.0	2614.0
	Other	37	2134.16	5548.41	534	3.0	32479.0
Turnover (2018 constant GBP)	Green	24	52460.43	162168.87	0.781009	5536.78	796730.77
	Brown	7.0	2383649.68	5809712.39	5535.84	226389.11	15555134.37
	Mixed	3	5899021	2666851.57	4200947.76	4523308.69	8972806.54
	Nuclear	1	771766.0132		771766.0132	771766.0132	771766.0132
	Other	37	880623.46	3151639.97	48.30091	59077.76	18347949.96

4.3.3 Innovation by UK OEMs & external drivers

This section discusses the trends highlighted in sections 4.1 to 4.3 in light of some of policy and market factors that could affect them. First, we note that the pattern of patent filing, both in aggregate and at the technology-level, continues to exhibit co-movement with public energy R&D spending. This is in line with, e.g., Johnstone et al. (2010); Dechezleprêtre and Glachant (2014), which find a positive effect of publicly funded R&D on patenting, and with the evidence of spillovers between academic research and some types of government R&D and the private sector.

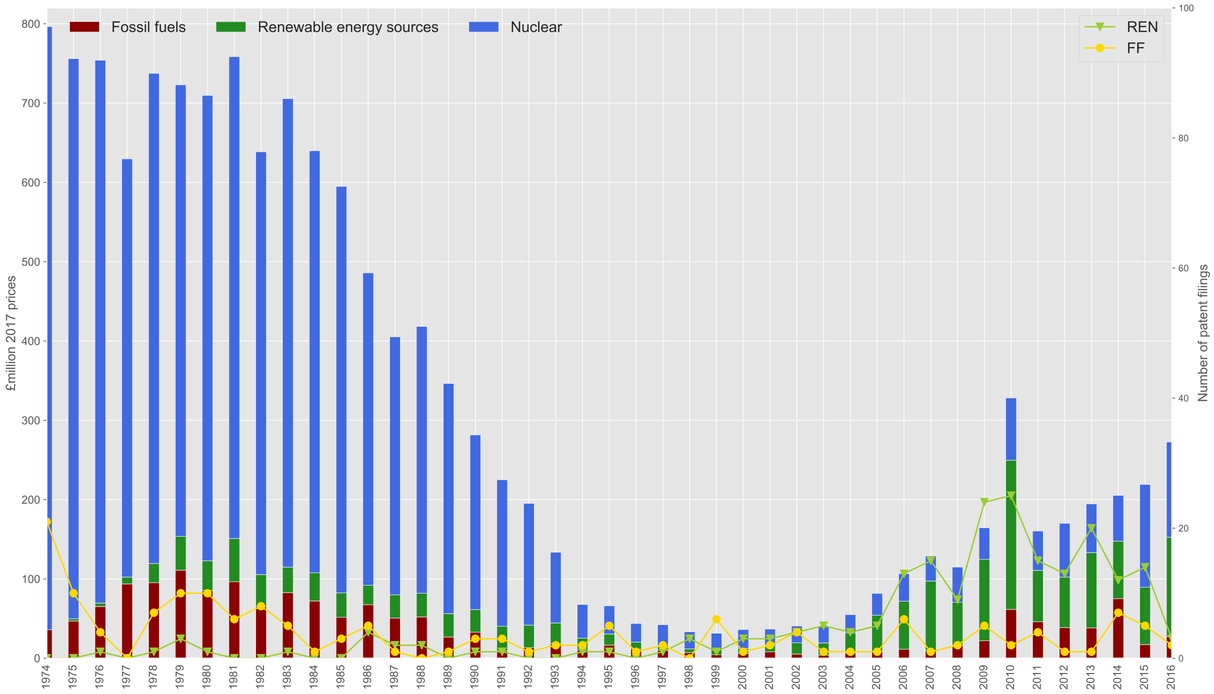

Figure 13: Public energy R&D spending and patent filings, by technology

In addition, as noted by (Popp, 2017), governments around the world continue to use energy R&D budgets as a key policy tools, not least as part of their climate change mitigation strategies. This line of work reminds us that, however central R&D by private institutions is to knowledge accumulation, innovation by and the role of other energy "research institutions" (e.g. national laboratories,. . .) in sustaining the industry's aggregate innovating activity cannot be ignored. In fact, "research not only funded but also performed by the government does appear to play an important translational role linking basic and applied research." (Popp (2017), p.1581).

Second, our review of the literature also pointed at the importance of the market environment for firms' innovation activity. In particular, we highlighted the relevance of (i) market incentives (e.g. oil prices) (ii) public policies (e.g. climate policy). Figure 14 shows patent filings in REN technologies together with a fossil fuel price index for the UK and the nominal EU ETS allowance price (in EUR/tCO_2e) taken from Dolphin et al. (2019). The co-movement is apparent.

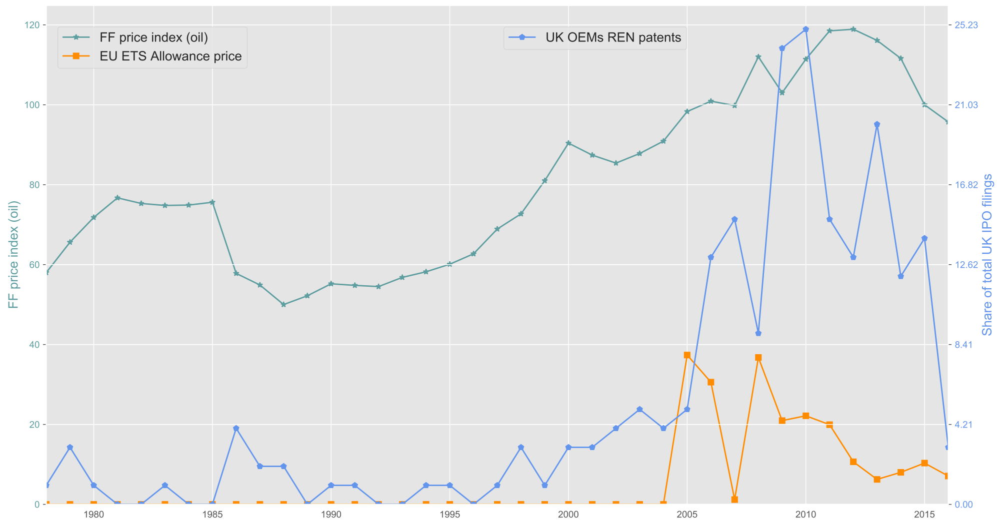

Figure 14: UK OEMs REN patent applications and fossil fuel & carbon prices (1978-2016)

5 Discussion and policy implications

The analysis presented above discussed the trends in patent filings in the UK Electricity Supply Industry over the period 1955-2016. This analysis identified the set of UK-based actors from which these filings originate and, in particular, shed a more precise light on innovation activity by upstream original equipment manufacturers. The trends identified in the sample of patents used in this study confirms the decline of innovation by downstream UK ESI actors and the shift of this activity toward upstream OEMs. The shift out of the UK of innovation by downstream actors following liberalisation and their passage into foreign ownership was already documented (Jamasb and Pollitt, 2011) but this study presents a further confirmation of this observation.

Second, the first part of the analysis also highlights the role of a few large public (e.g. UK AEA) or private (Rolls-Royce Plc) actors in the development of specific technologies. The UK AEA was instrumental in the development of nuclear electricity generation technologies whereas Rolls-Royce, building on its expertise in the design of jet engine turbines, played a crucial role in the development of fossil-fuel based electricity generation technologies. This highlights that large institutions have the potential to trigger innovation activity among a large set of actors across the industry, strengthening the case for (public) support for these institutions. Further evidence of this observation could be obtained by looking at the co-patenting activity of these actors.

Next, we observed that a conjunction of increased UK energy R&D spending, strengthened climate policy and high fossil fuel prices might have induced an acceleration of innovation activity between 2006 and 2010, especially by generating a high number of small new (technological) entrants in REN technologies. Indeed, a notable observation of this analysis was that innovation in renewable generation technologies has been brought about by new, small and technologically specialised firms; which might have been helped by the lower sunk cost to R&D in such technologies compared to fossil fuel or nuclear electricity generation technologies. The immediate policy implications of these observations is that in order to sustain innovation in these technologies, governments ought to (i) tailor policies in ways that specifically support (the growth of) young, small firms, (ii) keep barriers to entry low.

However, the revival in priority patent filings pertaining to these technologies does not seem to have been sustained, questioning whether the UK policy environment was appropriate to turn these firms into sources of sustained innovation.[34] Moreover, while a significant proportion of recent innovation activity was directed at renewable generation technologies, filing in fossil fuel generation technologies has continued, suggesting that little reallocation of R&D resources has taken place within firms. This should concern policy makers looking to make the power sector quickly transition to renewable generation technologies. Hence, improving our understanding of how to incentivise within firm resource (re)allocation constitutes an important research theme.

Finally, we reemphasise the scope of the study and point to additonal avenues for further research. First, the analysis was based on a sample of priority patent filings at

[34] We note, however, that part of this decrease in filing activity might be related the maturation of technologies, which in itself would be expected to lead to a decrease in the innovation and patent filing activity (Haupt et al., 2007; IEA, 2019).

the UK Intellectual Property Office. While this is in line with the objective of the study, it is important to note that these filings might not represent all filings by UK-based actors. Indeed, these entities may have filed (priority) patents either at other national patent offices, at the European Patent Office or at the World Intellectual Property Office. In that respect, searching for all priority patent filings by the actors identified in this study (based on their PATSTAT psn_{id}) could shed further light of the results discussed here. Second, we note that the set of patenting actors was identified based on a sample of patents retrieved through an ML and actors-based search. This approach implies that we have identified actors that have filed at least one patent over the period 1955-2016 but that we do not observe actors relevant to the ESI but that have not filed any patent over the period. In other words, our sample provides information about the intensive margin rather than the extensive patenting margin. Third, the discussion presented in this paper pertains specifically to the filing dynamics. As such, this study can't shed light on the value of the decline of institutions like the UK AEA or Electricity Council.

6 Conclusion

The world's commitment to keeping global average temperature increase below 2°C makes the further reduction of GHG emissions by the electricity supply industry in developed and developing economies alike an absolute necessity; even more so if the decarbonisation of other sectors of the economy is to be achieved by their 'electrification'. This will require further deployment of existing CO_2-abating technologies and the development of new ones. However, the latest patent filing data available suggests that innovation by UK-based actors has slowed. Moreover, among these declining filings, those related to fossil fuel generation technologies have picked up again. These trends must be reversed.

Given the predominant role that OEMs seem to have recently acquired, understanding the innovation patterns of these entities is of the essence. In this respect, the above analysis highlighted a few salient observations: (a) a majority of patents are filed by firms that are active in both fossil fuel and renewable electricity generation technologies, (b) but 'mixed' firms have filed significantly more fossil fuel generation technologies patents than renewables patents; and only during the period 2007-2013 have these firms filed more in the latter category than in the former (c) the increase in renewable generation technologies patent filings observed between 2005 and 2011 led to an increase in the number of technological entrants (i.e. firms patenting for the first time).

The evidence available so far shows that while prior policies have been successful at triggering technological entry, it has failed to create a (self-)sustained stream of innovation in "green" electricity supply technologies. Hence, the analysis suggests that any successful policy aiming at reversing the above trends ought to focus on supporting young, new entrants and turn them into sources of sustained innovation. This is especially important given that, at present, there is no large (UK-based) innovation actor in such technologies that could play a similar role as Rolls-Royce did for fossil fuel-based technologies.

Finally, we note that, historically, in the UK and other OECD economies, this innovation activity has originated from a variety of actors, ranging from government-owned vertically integrated utilities or research bodies to private entities, especially original equipment manufacturers and that innovation activity across this range of actors has been closely related to public authorities' strategic technological choices and energy R&D funding. Given that other electricity supply systems may, now or later, find themselves at a similar stage of their transition to a decarbonised electricity generation portfolio as the UK, its experience should be of particular interest to them.

Acknowledgements

The authors are grateful to Geert Boedt at the European Patent Office for helpful discussions about the patent system and PATSTAT online. The corresponding author gratefully acknowledges the support of the UK Economic and Social Research Council.

References

Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, 69(4):781–809.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and Innovation : An Inverted-U Relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

Aghion, P., Dechezleprêtre, A., Hémous, 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).

Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. In Nelson, R., editor, *The Rate and Direction of Inventive Activity: Economic and Social Factors*, volume 56, pages 191–209. Princeton University Press, Princeton, New Jersey.

BEIS (2019). Energy Trends 2019. Technical Report June, Department for Business, Energy and Industrial Stradegy.

Bound, J., Cummins, C., Griliches, Z., Hall, B. H., and Jaffe, A. B. (1984). *Who Does R & D and Who Patents ?* University of Chicago Press.

Bureau van Dijk (2018). Fame business structure database.

Calel, R. and Dechezlepretre, A. (2016). Environmental policy and directed technical change: evidence from the European carbon market. *The Review of Economics and Statistics*, 98(1):173–191.

Centre for Research & Development Monitoring (2017). ECOOM-EUROSTAT-EPO Person Augmented Table.

Crespi, F. and Scellato, G. (2015). Knowledge cumulability and path dependence in innovation persistence. In Antonelli, C. and Link, A., editors, *The Routledge Handbook of the Economics of Knowledge*. Routledge, London.

Dechezleprêtre, A. and Glachant, M. (2014). Does Foreign Environmental Policy Influence Domestic Innovation? Evidence from the Wind Industry. *Environmental and Resource Economics*, 58(3):391–413.

Dechezleprêtre, A., Glachant, M., Haščič, I., Johnstone, N., and Ménière, Y. (2011). Invention and transfer of climate change-mitigation technologies: A global analysis. *Review of Environmental Economics and Policy*, 5(1):109–130.

Dolphin, G., Pollitt, M., and Newbery, D. M. G. (2019). The Political Economy of Carbon Pricing: a panel analysis. *Oxford Economic Papers*.

European Patent Office (2013). Finding Sustainable Technologies in Patents. Technical report, European Patent Office.

European Patent Office (2018). Worldwide Statistical Patent Database (PATSTAT), version Spring 2018.

Eurostat (2018). Share of energy from renewable sources.

Gilbert, R. J. and Newbery, D. M. G. (1982). Preemptive Patenting and the Persistence of Monopoly. *American Economic Review*, 72(3):514–526.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4):1661–1707.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citations Data File: Lessons, Insights and Methodological tools. *Nber working paper series.*

Hall, B. H. and Mairesse, J. (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics*, 65(1):263–293.

Haupt, R., Kloyer, M., and Lange, M. (2007). Patent indicators for the technology life cycle development. *Research Policy*, 36(3):387–398.

Hicks, J. R. (1932). *The Theory of Wages.* Macmillan, London, UK, 1 edition.

IEA (2018). CO2 Emissions per kWh of electricity and heat output.

IEA (2019). Global patent applications for climate change mitigation technologies a key measure of innovation are trending down.

International Energy Agency (1999). Electricity Market Reform. Technical report.

Jaffe, A., Newell, R., and Stavins, R. (2002). Environmental policy and technological change. *Environmental and Resource Economics*, 22(May):41–69.

Jaffe, A. B. and Palmer, K. (1997). Environmental Regulation and Innovation : A Panel Data Study. *The Review of Economics and Statistics*, 79(4):610–619.

Jamasb, T., Nuttall, W. J., and Pollitt, M. (2008). The case for a new energy research, development and promotion policy for the UK. *Energy Policy*, 36(12):4610–4614.

Jamasb, T. and Pollitt, M. (2008). Liberalisation and R & D in network industries : The case of the electricity industry. *Research Policy*, 37:995–1008.

Jamasb, T. and Pollitt, M. G. (2011). Electricity sector liberalisation and innovation: An analysis of the UK's patenting activities. *Research Policy*, 40(2):309–324.

Jamasb, T. and Pollitt, M. G. (2015). Why and how to subsidise energy R+D: Lessons from the collapse and recovery of electricity innovation in the UK. *Energy Policy*, 83:197–205.

Johnstone, N., Haščič, I., and Popp, D. (2010). Renewable energy policies and technological innovation: Evidence based on patent counts. *Environmental and Resource Economics*, 45(1):133–155.

Joskow, P. L. (1998). Electricity sectors in transition. *Energy Journal*, 19(2):25–52.

Kamien, M. I. and Schwartz, N. L. (1975). Market Structure and Innovation : A Survey. *Journal of Economic Literature*, 13(1):1–37.

Klette, T. and Kortum, S. (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy*, 112(5):986–1018.

Kreuchauff, F. and Korzinov, V. (2017). A patent search strategy based on machine learning for the emerging field of service robotics. *Scientometrics*, 111(2):743–772.

Lanzi, E. (2010). Patent data analysis. *OECD Environment Working Papers*, (45):6–8.

Magerman, T., Van Looy, B., and Song, X. (2006). Data Production Methods for Harmonized Patent Statistics: Patentee Name Harmonization. Technical report.

Malerba, F. and Orsenigo, L. (1999). Technological entry, exit and survival: An empirical analysis of patent data. *Research Policy*, 28(6):643–660.

Mansfield, B. E. (1962). Entry , Gibrat's Law , Innovation , and the Growth of Firms. *American Economic Review*, 52(5):1023–1051.

Nemet, G. F. (2009). Demand-pull, technology-push, and government-led incentives for non-incremental technical change. *Research Policy*, 38(5):700–709.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (2006). The effects of economic and policy incentives on carbon mitigation technologies. *Energy Economics*, 28(5-6):563–578.

Nicolli, F. and Vona, F. (2019). Energy market liberalization and renewable energy policies in OECD countries. *Energy Policy*, 128(April 2018):853–867.

Noailly, J. and Smeets, R. (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management*, 72:15–37.

North Carolina Clean Energy Technology Center (2019). Database of State Incentives for Renewables & Efficiency (DSIRE).

Ofgem (2017). Electricity Network Innovation Allowance Governance Document.

OST (2014). Matching patent data with financial data. *Les notes pratiques de l'OST*, pages 1–5.

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, 92(1):160–180.

Popp, D. (2010). Innovation and Climate Policy. *National Bureau of Economics Research.*

Popp, D. (2017). From science to technology: The value of knowledge from different energy research institutions. *Research Policy*, 46(9):1580–1594.

Porter, M. E. and van der Linde, C. (1995). Toward a New Conception of the Environment-Competitiveness Relationship. *Journal of Economic Perspectives*, 9(4):97–118.

Sanyal, P. and Ghosh, S. (2013). Product market competition and upstream innovation: Evidence from the U.S. electricity market deregulation. *Review of Economics and Statistics*, 95(1):237–254.

Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy.* Harper, 3rd edition.

Torrisi, S., Hall, B. H., Bessen, J., Dernis, H., Macgarvie, M., Giuri, P., Grodal, S., and Mariani, M. (2010). Harmonizing and combining large datasets - an application to firm-level patent and accounting data.

A Patent data (and patents search)

Our main proxy for patenting activity in the UK is the number of patent applications contained in the EPO Worldwide Statistical Patents Database, version of Autumn 2018 (European Patent Office, 2018). We downloaded patents which had at least one IPC code starting with ‘B’,‘F’,‘G’, or ‘H’ for the years 1955-2017.[35] This represents 354760 patents.

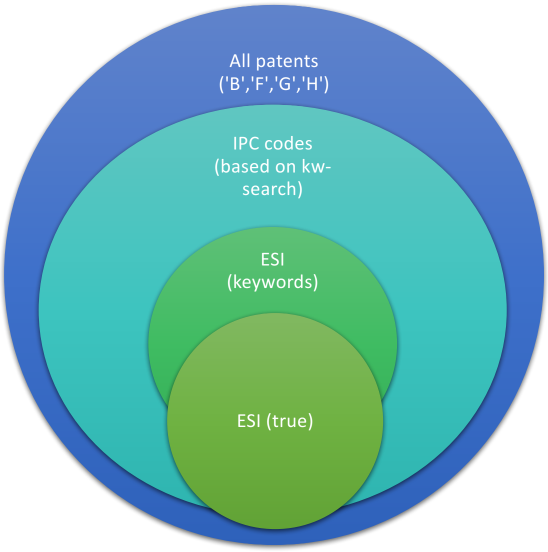

Figure 15: Patent sets

ML search To use the classifier on the said set, we need to create a training sample based on the text of patents related to electricity supply technologies and those related to other technologies. The approach taken to construct the sample and train the classifier builds on Kreuchauff and Korzinov (2017) and involves the following steps:

1. The classifier is trained on a sample of 240 patents, which includes 126 patents pertaining to electricity supply technologies and 116 patents pertaining to other technologies. This sample is constructed as follows. First, we randomly select 200 patents from the patent ensemble comprising all patents with at least one IPC code

[35] As mentioned earlier, we downloaded filings up to 2017 but truncated our sample to 2016 to account for a lag between filing and actual recording in the database. This covers all the IPC codes pertaining to fossil fuel and renewable electricity generation technologies, as identified in Lanzi (2010) and Johnstone et al. (2010) respectively, with the exception of the following codes: B01J8/20-22, C10J;C10L 5/40-48,C10L 1,C10L 3,C10L 5,E04D 13/18

in the list of those associated with the patents identified by the keywords-based search and 130 patents from the sample identified by the keywords search.[36] The titles and abstracts of all 330 patents were read so as to manually classify the them between (i) electricity supply technologies and (ii) other technologies. In the sample of patents drawn from the keywords-based ensemble, we identified 14 patents that were "false positives" whereas in the IPC ensemble we identified 22 "false negatives", which left us with 138 patents identified as belonging to the former category and 192 identified as belonging to the latter. In our training sample, we included all "false positives" of the keywords-based ensemble and 100 non electricity supply related patents of the IPC ensemble, as well as all "false negatives" of the IPC ensemble and all electricity supply related patents of the keywords-based ensemble. This, removing duplicates, left us with 114 patents pertaining to technologies unrelated to electricity supply and 126 patents pertaining to electricity supply technologies. These 240 patents constitute our training sample.

2. The text of the 240 patents titles and abstracts is prepared for classification

    (a) Structure the text data (application title <u>and</u> abstract). That is, for for each patent application: (i) merge patent title and abstract in one element and split into a single list of words; (ii) transform all string characters into lowercase characters; (iii) remove blank entries, stop words, empty spaces and numbers; (iv) extract the stem of each word; (v) Generate n-grams;

    (b) Derive normalised word and n-gram frequencies (across all patent applications)

    (c) Select features for classification. Not all features identified are carry meaningful information from a classification perspective. In other words, they add noise. Following Kreuchauff and Korzinov (2017), we kept only the features that appeared in at least 2% of the patent applications.

[36] The sample size is determined by the selection algorithm. We aimed for a sample size as close to 100 as possible, representative of the keyword search queries performed in proportion of the patents identified by each of them in the keywords ensemble, and under the constraint that at least one patent from each query.

3. Train our random forest classifier on a (training) sample. This comprised three iterative steps (common to almost any machine learning approach): training of the model, its evaluation, and optimisation. Finally, the classifier with the best model fit was applied to some test data;

Table 5: Classification report

	Precision	Recall	f1 score	No of patents in test set (support)
Non ESI	0.77	1	0.87	23
ESI	1	0.81	0.9	37
Avg./total	0.91	0.88	0.88	60

4. Apply trained classifier to full sample.

All steps describe above were performed using the python programming language and the following libraries: pandas (for data handling), nltk (for natural language processing), scikit-learn (for machine learning). Note that as a by-product of our "augmented" keywords-based search we also get a sample of patents selected only based on keywords.

Patents by NACE2 category The sample of identified entities spans a wide range of NACE classes.[37] 34% of the patents identified by our search strategy are associated with companies whose primary affiliation is the NACE "28.1 Manufacture of general-purpose machinery" or "28.11 Manufacture of engines and turbines, except aircraft, vehicle and cycle engines" class (and not the "32 Electricity, Gas, and Steam" class), followed by class "25.3 Manufacture of steam generators, except central heating hot water boilers" – , and class "26.5 Manufacture of instruments and appliances for measuring, testing, and navigation". This is a reflection of innovation activity taking place at the level of equipment manufacturers and illustrates the challenges that relying on NACE classes might pose for the definition of an industry.

[37]Graphical evidence is presented in figure 16 in appendix A.

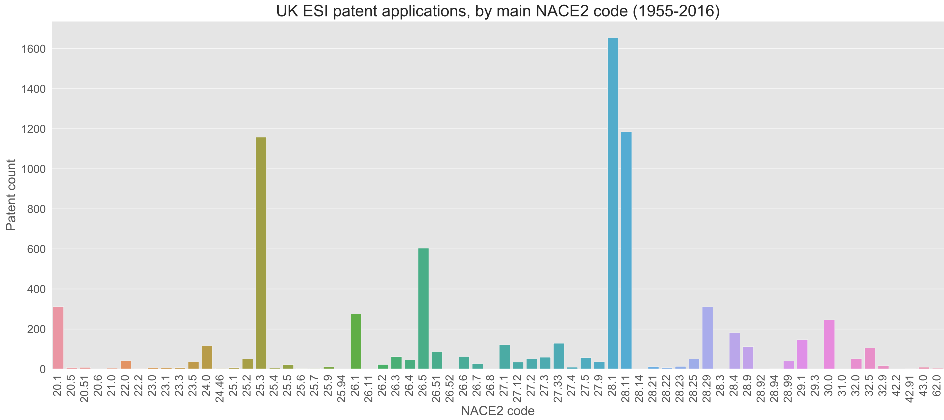

Figure 16: UK ESI patent applications, by main NACE2 class

This figure excludes filings by individuals. Including these filings changes the count of patents in each NACE category but does not change the proportional distribution across them. Within each class, counts are based on the weighted-average count, i.e. 1 over the number of classes with which each patent is associated.

List of actors List of actors used in our actors-based search and for which at least one patent filing entry was returned

Table 6: List of actors

BRITISH GAS TRADING
BRITISH NUCLEAR FUELS
CENTRAL ELECTRICITY GENERATING BOARD
CENTRAL ELECTRICITY GENERATING BOARD MARTIN R E
CENTRAL ELECTRICITY GENERATING BOARDS
CENTRICA CONNECTED HOME
CO-OPERATIVE ENERGY
DRAX POWER
EA TECHNOLOGY
ELECTRICITY COUNCIL
ELECTRICITY COUNCIL HODGETT D L FUNG H
ELECTRICITY COUNCIL ROBINSON G
ELECTRICITY COUNCIL THE
FLEXITRICITY
INNOGY
INNOGY TECHNOLOGY VENTURES
LONDON UNDERGROUND
MAGNOX ELECTRIC
NATIONAL GRID COMPANY
NATIONAL POWER
NORTH OF SCOTLAND HYDRO-ELECTRIC BOARD
NORTHERN IRELAND ELECTRICITY
NPOWER
POWERGEN
RWE INNOGY
SCOTTISH HYDRO-ELECTRIC
SCOTTISH NUCLEAR
SCOTTISH POWER
SOUTH OF SCOTLAND ELECTRICITY BOARD
UNITED KINGDOM ATOMIC ENERGY AUTHORITY

B Database queries

B.1 PATSTAT online

Below are the SQL queries used to query the PATSTAT database via the online platform PATSTAT online. PATSTAT contains bibliographical and legal status data originating from 90 patent issuing authorities, including regional patent offices like the EPO.

The first query is designed to retrieve all PATSTAT tables (except table *tls203_appln_abstr* containing the patent abstracts) pertaining to patent applications filed between 1955 and 2017 and with at least one technology code starting with H, G, F or B . In practice, the query is executed separately for five different time periods: 1955-1964, 1965-1973, 1974-1985, 1986-1996, 1997-2017. The start and end years of these periods are substituted for YEAR_START and YEAR_END in the code below.

The second query is designed to retrieve the abstracts of all patents identified by the first query.

```
-------------Retrieving Patent Applications and Titles----------------------
SELECT distinct app.[appln_id]
      /*,app.[appln_auth]
      ,app.[appln_nr]
      ,app.[appln_kind]
      ,app.[appln_filing_date]
      ,app.[appln_filing_year]
      ,app.[appln_nr_original]
      ,app.[ipr_type]
      ,app.[earliest_filing_id]
  ,dbo.GROUP_CONCAT_DS(DISTINCT ipc_class_Symbol ,N' ,' ,1) IPC */

FROM  [patstat2018a].[dbo].[tls201_appln] app
join tls203_appln_abstr on app.appln_id = tls203_appln_abstr.appln_id
```

```
join tls209_appln_ipc on app.appln_id = tls209_appln_ipc.appln_id
       and left(ipc_class_symbol, 1) in ('H','G','F','B')
where app.appln_auth = 'GB' and app.appln_filing_year between YEAR_START
        and YEAR_END
        and (app.earliest_filing_id = app.appln_id or app.appln_id in
        (select tls204_appln_prior.appln_id from tls204_appln_prior
        join tls201_appln
        as prior on tls204_appln_prior.prior_appln_id = prior.appln_id
        where appln_auth = 'GB'))
and app.appln_id < 900000000

order by app.appln_id desc

-------------Retrieving Patent Abstracts----------------------
SELECT distinct top 500000 app_abstr.[appln_id], app_abstr.[appln_abstract]
      /*,app.[appln_auth]
      ,app.[appln_nr]
      ,app.[appln_kind]
      ,app.[appln_filing_date]
      ,app.[appln_filing_year]
      ,app.[appln_nr_original]
      ,app.[ipr_type]
      ,app.[earliest_filing_id]
  ,dbo.GROUP_CONCAT_DS(DISTINCT ipc_class_Symbol ,N' ,' ,1) IPC */

FROM  [patstat2018a].[dbo].[tls203_appln_abstr] app_abstr
join tls201_appln on app_abstr.appln_id = tls201_appln.appln_id
join tls209_appln_ipc on app_abstr.appln_id = tls209_appln_ipc.appln_id
       and left(ipc_class_symbol, 1) in ('H','G','F','B')
```

```
where tls201_appln.appln_auth = 'GB'
and tls201_appln.appln_filing_year between 1955 and 2017
and (tls201_appln.earliest_filing_id = tls201_appln.appln_id
        or tls201_appln.appln_id in
       (select tls204_appln_prior.appln_id from tls204_appln_prior
       join tls201_appln  as prior on tls204_appln_prior.prior_appln_id =
       prior.appln_id
       where appln_auth = 'GB'))
and app_abstr.appln_id < 900000000

order by app_abstr.appln_id desc
```

B.2 FAME

The FAME database interface allows to perform company searches based on their Company Registration Number (CRN). In particular, it allows to upload a list of CRN on the interface and retrieve information on the associated companies. The extracted information includes: Global Ultimate Owner information (name, address, NACE Industrial classification code), number of employees, turnover, R-D expenditures, primary and secondary NACE code, date of incorporation.

C Technology codes

Table 7: IPC/CPC Technology codes

Technology	IPC class	CPC class	Source
Electricity generation			
Fossil Fuel			
	F01K		Lanzi (2010)
	F02C		Lanzi (2010)
	F02G		Lanzi (2010)
	F22		Lanzi (2010)
	F23		Lanzi (2010)
	F27		Lanzi (2010)
Efficiency-enhancing fossil fuel			
	F23C5/24		Lanzi (2010)
	F23C6		Lanzi (2010)
	F23B10		Lanzi (2010)
	F23B30		Lanzi (2010)
	F23B70		Lanzi (2010)
	F23B80		Lanzi (2010)
	F23D1		Lanzi (2010)
	F23D7		Lanzi (2010)
	F23D17		Lanzi (2010)
	B01J8/20-22		Lanzi (2010)
	B01J8/24-30		Lanzi (2010)
	F27B15		Lanzi (2010)
	F23C10		Lanzi (2010)
	F22B31		Lanzi (2010)
	F22B33/14-16		Lanzi (2010)
	F01K3		Lanzi (2010)
	F01K5		Lanzi (2010)
	F01K23		Lanzi (2010)
	F22G		Lanzi (2010)
	F02C7/08-105		Lanzi (2010)
	F02C7/12-143		Lanzi (2010)
	F02C7/30		Lanzi (2010)
	F01K23/02-10		Lanzi (2010)
	F02C3/20-36		Lanzi (2010)
	F02C6/10-12		Lanzi (2010)
	F02B1/12-14		Lanzi (2010)
	F02B3/06-10		Lanzi (2010)
	F02B7		Lanzi (2010)
	F02B11		Lanzi (2010)
	F02B13/02-04		Lanzi (2010)
	F02B49		Lanzi (2010)
	F01K17/06		Lanzi (2010)
	F01K27		Lanzi (2010)
	F02C6/18		Lanzi (2010)
	F02G5		Lanzi (2010)
	F25B27/02		Lanzi (2010)
Renewables		Y02E 10	European Patent Office (2013)
	F03D1-F03D11		Johnstone et al. (2010)
	F03G6		Johnstone et al. (2010)
	F24J2		Johnstone et al. (2010)
	H01L27/42		Johnstone et al. (2010)
	H01L31/04/78		Johnstone et al. (2010)
	H02N6		Johnstone et al. (2010)
	E04D13/18		Johnstone et al. (2010)
	F24J3		Johnstone et al. (2010)
	F03G4		Johnstone et al. (2010)
	F03G7/04		Johnstone et al. (2010)
	E02B9/08		Johnstone et al. (2010)
	F03B13/10-26		Johnstone et al. (2010)
	F03G7/05		Johnstone et al. (2010)
Nuclear		Y02E 30	European Patent Office (2013)
	G21F		Authors
	G21G		Authors
	G21K		Authors

Table 8: IPC/CPC Technology codes (cont.)

Technology	IPC class	CPC class	Source
Jet and gas engine turbines	B64C		Authors
	B64D		Authors
	F01D		Authors
	F02K		Authors
	F04D		Authors
Electricity transmission and distribution			
	H01B		Authors
	H01G		Authors
	H01H		Authors
	H01L 39		Authors
	H01T		Authors
	H01R		Authors
	H02G		Authors
Other technologies			
Instruments	G01		Authors
	G02		Authors
	G03		Authors
	G06		Authors
	G08		Authors
Energy storage	H02J		Authors
	H01M		Authors
Pollution control	B01D		Authors
Equipment manufacturing methods	B21C		Authors
	B22F		Authors
	B25J		Authors
	B29C		Authors
Other general engineering	F01D1		Authors
	F01D9		Authors
Other technology classes filed by UK AEA or EC	B01J		Authors
	H01F		Authors
			Authors
			Authors
			Authors
Other	F15C		Authors
	F16		Authors