Fuel prices and the invention crowding out effect: releasing the automotive industry from its dependence on fossil fuel

Nicolò Barbieri*
University of Ferrara

Abstract. The paper aims to shed some light on the impact of fuel prices and technological relatedness on green and non-green patenting dynamics and lock in to fossil fuel technologies. Specifically, we investigate whether green technology efforts come at the expense of other environmental or non-environmental invention activity. To do so, we employ Self-Organised Maps (SOMs) to detect the main technological domains exploited by the automotive industry during 1982-2008, using Triadic Patent Families (TPF) to proxy for the technological efforts in each technology field.

The paper adds to the literature by examining explicitly whether fuel prices (used as a proxy for carbon tax) and technological proximity foster the substitution of non-green patents by green ones. In addition, we provide a novel contribution by testing whether these determinants impact on the competition among low-emitting vehicles.

Our findings suggest that higher, tax-inclusive fuel prices are effective at redirecting patenting activity from non-green to green technology fields. Moreover, we observe that tax-inclusive fuel prices also induce a shift in patenting activity when we perform the analysis solely on green technology fields. Although this might result in potential lock-in to sub-optimal substituting technologies, our findings suggest that competition in the domain of environmental technology is focused mainly on 'greening' conventional cars and developing low-emission vehicles.

Keywords: environmental technologies; Self-Organising Maps; crowding out; fuel prices; patent data.

*Contact information: Department of Economics and Management, University of Ferrara, Via Voltapaletto, 11, 44121, Ferrara, Italy. E-mail: brbncl@unife.it or n.barbieri@unibo.it

1 Introduction

The impact of prices and policies on the development and adoption of clean technologies has been investigated extensively over the last decades. Although this body of research provides evidence that prices and policies affect environmental technological change (surveyed in Popp et al., 2010 and Barbieri et al., 2016), few studies focus on whether they affect the competition between clean and dirty inventive activities. Indeed, although alternative technological trajectories might improve environmental performance, evolutionary economists emphasise that the process of technology selection is path dependent, not predictable ex ante and irreversible; thus, the market may select sub-optimal technologies based on the increasing returns to adoption (Arthur, 1989; Bruckner et al., 1996; Frenken et al., 2004). This conservatism in market selection, on the one hand, has a negative effect on the probability that alternative technologies will be adopted ('self-reinforcement') and, on the other, allows producers to take advantage of economies of scale and R&D investments (David, 1985).[1] In addition to path dependence in technology adoption, Acemoglu et al. (2012) states that the type of innovation produced follows a path-dependent process, which provides incentives for firms that have previously developed dirty technologies to continue to develop them in the future.

It should be noted also that the evolutionary process at the basis of technological change highlights that the success of technological advancement cannot be determined ex ante (Nelson and Winter, 1982). This is due mainly to the uncertainty surrounding the design and planning processes. For example, successful technological advances are the result of a process in which, at any time, a range of technological opportunities is undertaken and proposed to the selection environment (Gelijns et al., 2001). Therefore, there is competition among innovations, and the prevailing technology is determined by ex post selection (Gelijns et al., 2001).

Furthermore, in the case examined in the present paper, that is, the automotive industry, we find that technological uncertainty also affects the development of low-emissions vehicles. One source of

[1]In David (1985), the author ascribes the lock in to QWERTY to technical interrelatedness, economies of scale and the quasi-irreversibility of investment.

uncertainty is linked to the capability of alternative cars to substitute for conventional vehicle designs and another is related mainly to competition *between* alternative vehicles since, in the current climate, it is unclear which alternative option should be preferred from both an economic and environmental perspective (Frenken et al., 2004).

Against this complex background, where uncertainty, path-dependence and competition prevail, several authors highlight that policy intervention is one of the main factors that might allow socio-technical lock-ins to be reduced (Faber and Frenken, 2009; Rennings et al., 2013) and, specifically, for lock in to Internal Combustion Engine Vehicles (ICEV) be avoided (Cowan and Hultén, 1996).[2] Recently, several authors have highlighted the role of environmental policies in inducing development of environmentally-sound technologies (Popp et al., 2010; Bergek et al., 2014). However, although it has been shown that environmental policies lead to increasing innovative performances and market competitiveness (Porter and van der Linde, 1995), the production of eco-innovations can have secondary effects such as environmental rebound, the green paradox and crowding out (van den Bergh, 2013). Indeed, environmental policies can lead to higher opportunity costs derived from real resources (financial and human) requirements for the development and adoption of alternative technologies to comply with policy objectives (Jaffe et al., 2002). They can trigger innovation in green technology domains which may drive inventive activity from non-environmentally friendly to green innovations, becoming a potential source of innovation crowding out.

The present paper investigates the role played by fuel prices (our proxy for a carbon tax) on technology dynamics, in a sample of automotive firms, i.e. we analyse the effectiveness of fuel prices for breaking the link with ICEV technologies. In this context, the crowding out effect, generated by higher fuel prices, may favour this objective; although crowding out of any type of innovation reduces the social benefits[3] and decreases competition, it might help to delink the automotive industry from its

[2]In addition to regulation, the authors identify other factors such as problems in existing technology, radical technologies, differences in taste, niche markets and scientific achievements (Cowan and Hultén, 1996).
[3]The social returns from research are greater than the private returns to firms (Mansfield et al., 1977; Pakes, 1985; Jaffe, 1986).

dependence on fossil fuel, that is, it might work to reduce innovation activity in ICEVs in favour of Low-Emissions Vehicles (LEVs).

With a few exceptions which are discussed below, this topic has been mostly unexplored, and very little debate has been about policy-driven crowding out effects and, especially, 'what' is being crowded out. If improvements in technologies with negative environmental effects are crowded out to favour advances in green technologies, the costs to society of this crowding out could be reduced (Popp, 2005) or, in the case that this crowding out affects other environmental technology efforts, increased. Therefore, we test whether invention activity in the area of environmental technologies comes at the expense of other green inventions.

The paper is structured as follows: Section 2 discusses the related literature; Section 3 explores the main features of the automotive technological system, presents the data and identifies the main technological trajectories using Self-Organising Maps (SOMs). Section 4 describes the building of the main variables and the empirical model and Section 5 discusses the results. Section 6 concludes the paper.

2 Literature review

In a recent review of studies investigating eco-innovation from an evolutionary perspective, Cecere et al. (2014) emphasise that technological, social, organisational and institutional lock-ins affect environmental innovation development and adoption. This suggests that firm-level strategies, technological niches and regulation are the key to overcoming path dependence on the dominant technological designs. In particular, there is one literature stream that provides evidence of the effectiveness of environmental policy for boosting green technologies (surveyed in Popp et al., 2010) and sheds light on its potential to unlock technological systems. Several studies of environmental regulation assess whether environmental policy fosters technological change towards a more sustainable path. However, this body of work on the policy-induced development of environmental technologies does not provide direct insights into the potential competition between non-green and green inventions. To understand the overall effect of green policy on the economic system, it is necessary to take account of potential secondary consequences of policy implementation in order to appreciate their overall impact, beyond the development of new green technological efforts. Indeed, environmental inventions may come at the expense of non-green innovations or may become complements in firms' innovation portfolios. However, analysis of the crowding out effect has been hampered by the difficulties involved in addressing this issue empirically. It is difficult also, even ex post, to identify whether a change of direction in innovation activity is due to policy intervention or research opportunities and firm strategies.

A seminal work that discusses the presence of a crowding out effect is Gray and Shadbegian (1998). The authors examine the impact of environmental regulation stringency in the pulp and paper industry. In their study, crowding out affects decisions about investment in pollution abatement, and productive (non-environmental) capital investments. Their results suggest that investment in pollution abatement crowds out other productive (non-abatement) investments within a plant.

Marin (2014) uses a dataset of Italian manufacturing firms and provides evidence (at least in the short run) that environmental innovation comes at the expense of non-environmental innovation. This possible crowding out is driven mainly by the lower returns which distinguish eco-innovation from other investments, coupled with the constrained financial resources devoted to R&D activities.

When firms are not financially constrained, a decrease in non-environmental innovations caused by an increase in green innovation, does not always imply that the crowding out effect reduces the social and private benefits. Popp and Newell (2012) investigate whether the increase in environmental R&D spending leads to lower levels of R&D investment in other fields. First, the authors find no evidence of crowding out across sectors, 'mitigating the concern that new energy R&D programs will draw resources away from other innovative sectors of the economy' (Popp and Newell, 2012, p. 990). Second, using patent data to proxy for R&D expenditure, they examine whether this hypothesis holds within sectors and find that an increase in alternative energy patents leads to a decrease in other patents. However, the absence of financial constraints in the firms studied might suggest that the crowding out effect is driven by changes in market opportunities. This second result underlines the positive environmental effect of crowding out, which seems to induce development of green technologies at the expense of dirty ones, which helps to satisfy environmental policy objectives.

More evidence of an R&D offsetting comes from Kneller and Manderson (2012). Their results highlight that an increase in environmental compliance costs boosts environmental innovation, although the effect of environmental expenditures does not have a positive impact on the total amount of R&D investment, suggesting that environmental R&D crowds out non-environmental R&D[4].

As a result mainly of the research questions addressed, most of these studies do not directly examine the role of environmental policies in this framework. An exception is Hottenrott and Rexhäuser (2013), which employs survey-based data to identify which firms introduce environmental technologies as a consequence of policy compliance behaviour. Their study suggests that, while there is evidence that environmental innovation crowds out firms' in-house R&D expenditure, this does not seem to

[4] The authors highlight the lack of evidence that environmental capital crowds out non-environmental capital.

influence the number of existing R&D projects, their outcome or the amount of investment in fixed assets (both innovation-related and other). In addition, the authors suggest that firms prefer to scale down long-term oriented R&D activities which are not connected directly to production and which provide relatively uncertain returns. A recent work by Noailly and Smeets (2015) on directed technical change, investigates the extent to which renewable energy patents can replace fossil fuel patents in the electricity generation sector, causing a shift in innovation activities. Their findings suggest that entry of specialised renewable energy firms reduces the fossil fuel-renewable energy technology gap, whereas fossil fuel prices, market size and knowledge stock increase the technology gap by stimulating the production of dirty inventions in mixed firms which innovate in both dirty and renewable energy technologies.

2.1 Research questions

Recent studies on directed technical change investigate the factors that might influence a shift towards environmentally friendly technologies (Acemoglu et al., 2012; Di Maria and Van der Werf, 2008). Following the theoretical framework proposed by Acemoglu et al. (2012), Aghion et al. (2012) empirically test the impact of fuel prices on patenting activities in the automotive sector. They emphasise that tax-inclusive fuel prices and firms' past knowledge stock influence the direction of technological change, inducing firms to invest more in alternative than in conventional vehicle technologies.

Our paper draws on these findings to analyse whether fuel prices encourage a shift from non-environmental to environmental inventions. Since fuel prices directly affect the returns from innovation and the firm's incentive to innovate, we employ fuel prices to measure the economic incentive to increase energy efficiency and develop alternatives to conventional vehicle technologies. Indeed, it has been acknowledged that fuel prices also have a negative effect on demand for vehicles (Busse et al., 2009; Allcott and Wozny, 2014). The first research question we address is:

1. Do tax-inclusive fuel prices induce a shift from non-environmental inventions to environmentally friendly invention activity?

There are two main propositions in the literature on technological substitution (David, 1985; Arthur, 1989). First, even when substituting technologies are available and are superior to the dominant technology, the presence of increasing returns to adoption means that technological substitution cannot be assured. Second, in a process of technological substitution, a pool of new technologies competes for dominance although lock-in to sub-optimal substituting technologies remains a possibility due to path dependence in sequential adoption decisions. Both propositions apply, at least in part, to the automotive industry due to competition between conventional and low-emission vehicles, and between alternative, environmental-friendly vehicles (Frenken et al., 2004).

Since the potential shift from non-green to green inventions might also affect the environmental domain due to the competition between low-emissions vehicle technologies (i.e. green inventions) at the expense of other green inventions, we investigate 'what' it is that is being crowded out. In this case, if fuel prices induce a resources shift from environmental technological domains to other green inventions, this will increase the risk of technological lock-in to a sub-optimal substituting technology due to the absence of a superior alternative technology from both an economic and environmental perspective. This leads to the second research question:

2. Do tax-inclusive fuel prices alter competition between alternative low-emissions vehicles? Does it cause a shift among environmental inventive activities?

3 The automotive technological system

3.1 Patent data in the automotive sector

To address our research questions, we focus on large incumbent automotive firms. Our choice has several motivations. First, due to the high impact on local and global air pollution, policy makers across the world are highlighting the need to decrease polluting emissions from vehicles. Over the last two decades several regulatory policies have been introduced to mitigate the impact on the environment of the transport sector. Second, numerous scholars have highlighted the presence of carbon lock-in in the automotive industry (Cowan and Hultén, 1996; Frenken et al., 2004; Aghion et al., 2012). Third, the automotive industry has been challenged by deep structural changes and financial uncertainties, which have led to reconsideration of the management of knowledge capital (Laperche et al., 2011). In addition to the industry dynamics (R&D rationalisation, R&D collaboration, etc.), increased demand for low-emissions vehicles combined with environmental regulation, has provided the incentive to develop new environmentally-sound technologies and reduce vehicle emissions levels. Fourth, intellectual property (particularly patent protection), especially in the automotive industry, plays a pivotal role in triggering profits and competitive advantage (Laperche et al., 2011).

Since our aim is to explore the dynamics of the inventive efforts in different technological fields, we employ patents as a proxy for invention. Griliches (1990) points out that patents, sorted by priority year, show a strong correlation with R&D expenditure. Also, patents are the only source of data on the technical features of invention activity, which is needed to test our hypotheses. However, patent data have some limitations (see, e.g., Griliches, 1990). First, inventions can be protected in other ways, such as lead time, industrial secrecy or purposefully complex specifications, than patenting (Frietsch and Schmoch, 2006). Second, the propensity to patent can vary depending on the countries in which they are filed, the type of technology and the risk of imitation (Cohen et al., 2000). Third, patents are heterogeneous with respect to their technical and economic significance. Fourth, there can be problems related to the variability in the quality of patent data (Lanjouw et al., 1998) and the selection process employed (keyword search, patent classification search, etc).

To take account of these limitations, we employ a methodology based on TPF, defined by the OECD as a 'set of patents taken at the EPO, USTPO and JPO that share one or more properties' (Dernis and Khan, 2004, pp.17). One such property is that patents must pertain to the same patent family.[5] Thus, we focus on high quality patent data since the most important inventions are protected in these three patent offices that make efforts to reduce the influence of the heterogeneity of patent office regulation systems (Dernis and Khan, 2004). To deal with patent sample selection problems deriving from the type of search conducted,[6] we collected the automotive firms included in the R&D scoreboards (IRI) from the 2006-2011 editions.[7] This results in a focus on firms involved in continuous, high levels of R&D investment. Incumbent firms are expected to be involved in large R&D programmes based on their consolidated financial and R&D capabilities (Cohen and Klepper, 1996). To obtain a picture of the corporate structures and standardised applicant names, we gathered patents filed by those 71 firms, retrieving their names from the Derwent Corporate Tree.[8]. This provided data on the complex globalised structure of the automotive industry, and reduced noise caused mainly by different spellings of assignee names.

[5] Patent families are defined by the OECD as "the set of patents (or applications) filed in several countries which are related to each other by one or several common priority filings" (OECD, 2009, pp.71).

[6] Patent data can be collected in various ways, but there are drawbacks to data collection using patent classification (Costantini et al., 2013) and applicant name (Thoma et al., 2010) searches.

[7] The number of firms ranked on the R&D Scoreboards before 2006 and after 2011 were very different (500 and 2,000 as opposed to 1,000 firms). Thus, the 2006-2011 rankings are homogeneous and comparable.

[8] The Corporate Tree tool covers the top 2,500 patenting companies for those authorities, and takes account of mergers, acquisitions, divestitures and spelling differences (but not reassignments). 6 firms were not included in the Corporate Tree; we found the relevant patents by searching on the applicant name field in the OECD Harmonised Applicants' Names database.

3.2 The automotive technology space

Using the Thomson Innovation database, we collected all the patent family applications filed by the sample of firms and obtained a total of 247,510 patent families, 54,370 of which were TPF. We distinguished between green and non-green TPFs by exploring their technological classification codes. Different technological classifications have been proposed to analyse the technological content of patent data. In this paper we use Cooperative Patent Classification (CPC) codes,[9] which provide a hierarchical and language independent classification of the patents' technical domains. This classification is particularly appealing for our study because it allows detection of green patents via the Y02 class "Technologies or applications for mitigation or adaptation against climate change". This allowed us to identify a subsample of environmental inventions in the dataset. The last row in Table 1 reports the total number of patents on which the analysis is performed, and the percentage of green patents (i.e., patents to which Y02 CPC codes are assigned) in the whole sample of patents.

Figure 1 depicts trends in green and non-green TPF applications sorted by earliest priority year. The histograms show that the percentage of green TPF per year rose steadily from 1980 to 2006 (when it peaked), after which it fell to 2009, and the opposite trend for the percentage of non-green patents. We observe also that the percentage of green and non-green TPFs in the total (respectively, green and non-green) number of TPFs over the whole period, increased sharply after 1990. However, while the percentage of non-green patents has fluctuated since 2000, green patents continued to increase up to 2006. These observations highlight that the growth in the number of green patents in recent years was due to environmental policy efforts to 'green' ICEV technologies and to develop new alternative vehicle propulsion systems (Barbieri, 2015).

However, in order to investigate which inventive activities affect the dynamics of technological effort, we distinguish also among the types of technologies included in the green and non-green technological fields. We take the share of CPC classes among inventions as a proxy for their technological similarity,

[9] CPC is a new classification introduced in the USPTO and EPO, which includes a section for emerging technologies (http://www.cooperativepatentclassification.org). For an application of CPC patent maps, see (Leydesdorff et al., 2015).

that is, the higher the number of the same CPC classes , the greater the technological relatedness of the patents. In contrast to approaches that use patents to measure relatedness between technological fields (Jaffe, 1986; Nesta and Saviotti, 2005; Breschi et al., 2003; to cite a few), we employ technological fields to map inventions based on their technological similarity.[10]

We constructed a distance-based patent map using the SOM technique (Kohonen, 1990; 2001). The SOM is a unsupervised neural network, which represents multidimensional data in a two-dimensional space, illustrating the similarity among input data. We employ this technique to map patents based on their technological similarity, using CPC codes assigned to each patent as input data.

3.3 Self-Organising Maps

3.3.1 How does SOM work?

SOMs are a nonlinear mapping projection technique, which identify similarities among input data and represent them in a network of interconnected nodes (neurons) to which the input items are assigned according to the Euclidean Distance (ED) between the nodes' weight vectors and input vectors. The SOM is based on two main processes: initialisation and training. The former refers to the assignment of weights to an empty grid of nodes (neurons) using two alternative techniques: randomised and linear initialisation. In the case of linear initialisation, initial node values are "selected as a regular, two-dimensional sequence of vectors taken along a hyperplane spanned by the two largest principal components of x" (Kohonen, 2013, p. 6), where x is a set of vectors representing the input data.[11] Due to its higher efficiency in the convergence of the training algorithm, linear initialisation techniques are preferred to random initialisation, which assigns node values randomly (Kohonen, 2013).

[10] Other works focus on patents to "link" technological fields, i.e., the same technological fields in two patents is evidence of the relatedness between fields. In our study, we consider similar technological fields as linking patents, i.e. the higher the number of common classification codes, the greater the similarity of the technological content of the patents.

[11] Note that the vectors associated to each observation, capture the multidimensionality of the input data.

During the initialisation phase, a vector of the initial values $m_i = [m_{i1}\ m_{i2} \dots m_{id}]$ (where d indicates the dimension of input vectors) is assigned to each neuron i. These vectors are of the same length as of the input vectors x. In the next step, for each input vector the training phase identifies the Best Matching Unit (BMU) (m_c) in the initialised map. The BMU is the neuron with the smallest ED from x, that is:

$$c = argmin\,\{\ \|x(t) - \ m_i(t)\|\}$$

where t represents a step in the sequence in which the sample of n-dimensional Euclidean vectors x is presented to the map (denoted by $x(t)$). Thus, in step t the neuron $m_c(t)$ is defined as BMU, meaning that it is the best match with $x(t)$ – compared to the other neurons in the grid. Once all the input vectors x are assigned to their BMU on the map, the training phase adjusts the initial values (m_i) of each neuron according to the set of input vectors x, assigned in the previous step. There are two algorithms - batch and stepwise recursive training - that can be used for the training step. We chose the former because it produces a final output independent of the order in which the input data are presented to the map's nodes. It also produces a steady solution more safely and more quickly than does the recursive stepwise training algorithm (Kohonen, 2013).

This process allows the input vectors to modify the weights of BMU (to which they are assigned) and the weights of the nodes within its neighbourhood set (N_i) – the set of nodes located within a certain radius from the BMU. The weight of each neuron on the map is updated by calculating the mean of the input vectors assigned to the neurons defined by the neighbourhood function. Neurons weights are calculated as follow:

$$m_i(t+1) = \frac{\sum_{j=1}^{n} h_{ic(j)}(t)x_j}{\sum_{j=1}^{n} h_{ic(j)}(t)}$$

where $c(j)$ is the BMU of a sample vector $x(j)$, n is the number of input vectors and h_{ic} is the Gaussian neighbourhood function defined as follow:

$$h_{ic} = \alpha(t)\ exp\left(-\frac{\|r_i - r_c\|^2}{2\vartheta^2(t)}\right)$$

where $0 < \alpha(t) < 1$ represents the learning-rate factor, which decreases as long as t increases; r_i and r_c are the vectoral locations on the map of the nodes i and c; and $\vartheta(t)$ is a monotonically decreasing function of t, which indicates the width of the neighbourhood kernel (Kohonen, 1998).

To obtain the SOM using the batch training algorithm involves the following steps:

1. Initialising the map's node weight vectors through linear initialisation;
2. Selecting the input vector x;
3. Traversing each node in the map:
    3.1. calculating the ED between x and the map's node weight vectors m_i;
    3.2. identifying the BMU (the node with the smallest distance) (m_c) and assigning x to that neuron;
4. Repeating steps 2 and 3 until all input vectors are assigned;
5. Computing the means of the input vectors assigned to each neuron and its neighbours which constitute the neighbourhood N_i (defined through the neighbourhood function);
6. Replacing the old values of m_i by the respective means obtained from 5;
7. Incrementing t and repeating steps 2 to 7 until the solutions can be regarded as steady.

Note that, in each iteration (from 2 to 7) the neuron weights are modified to resemble the input vectors. Therefore, by calculating the ED between each pair of neurons, we observe that the relative positions of map's nodes change, which captures the multidimensionality of the input data in an ordered and relative way. Thus, the distance between the nodes in the final output map can be used to proxy for their similarity. Similar (different) input data are positioned closer than (further away from) different (similar) data.

3.3.2 Graphical interpretation of the SOM algorithm

Figure 2(a) depicts how the input data were introduced in the present exercise. Each row represents a patent, the columns denote the CPC codes and the matrix values indicate whether a CPC is assigned to the patent (1) or not (0). Following initialisation, where weights are assigned to an empty map (Figure 2b), the batch training algorithm (Kohonen, 2013) is implemented. In each step, the SOM randomly selects an input (in our case a patent) and detects the map node with the lowest ED (BMU) between the input vector and the initialised nodes in the maps (Figure 2c). This step is repeated until all the input vectors are assigned to the map's nodes. A radius defines the neighbours for each BMU, that is, the set of nodes closest to the BMU (Figure 2d). Then the neighbour node weights are modified to make them more similar to the BMU, decreasing the ED between similar map units and increasing the ED between different units.

Use of the SOM provides numerous advantages. We are able to: i) locate patents in a technology space which indicates their similarity (the more similar their technological content, the closer they are located on the map); ii) define patent clusters which refer to the same vehicle component (e.g., hybrid engine, catalytic converter, battery,; brakes, etc.); and iii) measure the relatedness between these clusters. In comparison to other techniques and methodologies to retrieve the cognitive distance between technological fields, the SOM provides a distance-based output which locates patents according to their global and local similarity.

3.4 Exploring the technological space

Using the occurrence of patent classification codes assigned to each patent (CPC codes) as SOM input data allows us to capture the technological similarity underlying the patent dataset. In order to define technological clusters, we apply the non-hierarchical k-means clustering technique (MacQueen, 1967) on the SOM output. The objective of coupling SOM with k-means is to identify groups of patents that are technologically related. The properties of the SOM discussed above mean that each patent is mapped closer to (distant from) other similar (different) patents. We assume that when the same set of CPC codes is assigned to two patents their technological relatedness is at a maximum.

The SOM and k-means algorithm outputs are depicted in Figure 3(a), which shows the distance between nodes and their closest neighbours, that is, the Unified-distance Matrix (UMAT); Figure 3(b) depicts the 31 clusters identified by the k-means clustering process applied to the SOM.[12]

Table 1 lists the main keywords[13] associated with the 31 clusters of inventions identified by the joint procedure. Columns 3 to 4 in Table 2 show the number of patents and the percentage of green patents in each cluster. Note that the clustering exercise correctly identifies and positions green inventions, creating clusters consisting almost entirely of environmental patents.

Figure 3(b) shows also that green inventions are located at the bottom of the technology space. The main trajectories of technological advancement in the automotive industry are in this area of the technology space. Fundamental changes that occurred in the 1970s affected the car market. Growing concern over traffic congestion and air pollution accompanied by the oil crises, contributed to modifying the economic and social factors governing technological developments in that industry. Since that time, there have been several different technological trajectories including increased variety

[12] K-means is run multiple times for each k. The process selects the best alternative with regard to the sum of the squared errors. The Davies-Bouldin (Davies and Bouldin, 1979) index is calculated for each alternative.

[13] We collected the title and abstract of each patent per cluster and then examined the text in these groups of words using text mining techniques. After a cleaning process, which deleted the stopwords (e.g. a, the, then, if, etc.) and reduced the words to their stem (e.g. stemming becomes stem, automobile becomes automobil, etc.), we weighted each word using the Term Frequency/Inverse Document Frequency (TF/IDF). Finally, we ranked the weighted words in each cluster and chose the most representative of the first 20 words.

of LEVs competing with ICEVs, that is, Electric Vehicles (EV), Hybrid Vehicles (HV) and Fuel Cell Vehicles (FCV).

Moving from left to right in the bottom half of Figure 3(b), we observe the variety of LEV technologies that have influenced the main technological trajectories in alternative vehicles. On the left, Clusters 6 and 12 represent HV technologies integrating the ICE and the electric motor (Dijk and Yarime, 2010). HV technology was seen as promising, at least in the short run, for a transition from ICEV to FCV (Oltra and Saint Jean, 2009). The technology to the right of these two clusters is characterised by batteries implemented in HVs and EVs. Specifically, the inventions in Clusters 9, 10 and 22 exploit alternative battery systems, which are the main barrier to a sizeable electric car market. Technological variety in LEVs is completed in Cluster 20, which includes FCV. Finally, Clusters 2, 8, 27 and 31 include technologies which reduce the impact of ICEVs such as catalytic converters, turbochargers, direct injection, etc. These technological improvements are representative of what we described previously as the 'greening' of the dominant design in the automotive industry.

The remaining and the majority of the technology space, is characterised by non-green inventions. In the centre of the map, Cluster 11 has the highest share of nodes among all the other clusters, confirmed by it representing almost one-third of all automotive patents. This cluster includes heterogeneous components, such as mechanical and electronic apparatus (e.g., air conditioning systems, automatic door opening, etc.), and car designs not directly related to the powertrain system. We can identify two main technological effort trajectories which advance towards the upper part of the map. Clusters 4, 5, 25, 19, 17 and 1 include patents related to engine mechanical components and catalytic converters. The former refers to powertrain system technologies and the latter to end-of-pipe technologies outside the realm of green technologies (e.g., silencers). The left side of the map is characterised first by inventions related to battery systems (Cluster 30 and 7) and elements such as cruise assistance (Cluster 26) and control systems (Cluster 16). Clusters 13, 14, 15, 21 and 24 need separate discussion. These clusters include mechanical developments in transmission (13-15), suspension (24) and braking systems (21). Finally, Cluster 23 includes safety technologies and Cluster 3 includes tyres and pneumatics patents.

In order to test the effectiveness of this clustering exercise we collected all the patent citations in the patent dataset. Patents include citations to prior inventions, which constitute prior art. If patent A cites patent B, there is a technical relation between the two based on the knowledge included in previous patent (B) which the more recent patent (A) builds on (Martinelli and Nomaler, 2014). Figure 4 shows that the majority of patents cite inventions in the same cluster, which supports the reliability of the cluster analysis.

4 Variables and empirical model

This section presents the variables and the empirical model employed to study what influences automotive technological systems.

4.1 Dependent variable

The dependent variable, CO, allows us to measure the shift in inventive efforts made in each technological field. It is calculated as follows:

$$\Delta PAT_{z,t} = ma_{z,t} - ma_{z,t-1}$$

$$CO_{i,t} = \Delta PAT_{g,t} - \Delta PAT_{ng,t}$$

where ma is the patent count moving average,[14] z refers to specific clusters (listed in Table 1) and PAT is the growth rate of the ma for each cluster. The variable CO is the difference between the growth rate of the patent count moving average in a green cluster (g) and the same growth rate in relation to a non-green cluster (ng); i represents each pair of green and non-green clusters. Therefore, when the CO variable is positive, the green cluster growth rate is higher than the rate of growth in a non-green cluster. We assume that positive values of this variable imply a shift in technological advances towards more sustainable technologies. Similarly, the CO variable can be used to test the potential crowding out effect among green clusters, as follows:

$$CO_{s,t} = \left|\Delta PAT_{g1,t} - \Delta PAT_{g2,t}\right|$$

[14] We tested different year moving average spans, e.g. 3,4 and 5 years.

where CO is equal to the difference between two patent count moving averages related to $g1$ and $g2$ (with $g1 \neq g2$) with s representing each couple of green clusters. The absolute value helps the interpretation of the results since the output in this case is bidirectional.[15]

The strength of this approach resides in its ability to account for a relative increase (decrease) in patent counts related to both technological clusters (g vs. ng and $g1$ vs. $g2$), that is, a technological field increases more than proportionally to another.

4.2 Independent variables

4.2.1 Tax-inclusive fuel prices

We test whether the dependent variables defined previously are affected by fuel prices. A large body of literature has analysed the effect of fuel prices on innovation (see Crabb and Johnson, 2007; Hascic et al., 2009; Aghion et al., 2012, among others). These studies propose a consolidated framework which provides evidence of the positive impact of environmental policy on environmental innovation. In addition, if this variable has a positive impact on our dependent variable this suggests that higher fuel prices increase the probability that green inventions come at the expense of non-green invention activity and, thus, that rather than being *additional*, green technological efforts crowd out non-green efforts. In this case, we would suggest that environmental policies could be effective in reducing path dependence on conventional non-environmental technologies. On the other hand, a negative effect might suggest that even if environmental regulation induces firms to increase their green invention activity, it does not affect non-green technological improvements, which demonstrates their ineffectiveness at redirecting technological advances away from ICEV technologies.

[15] The patent count in each green technological fields is compared with that in every other green cluster except the focal one.

Following Aghion et al. (2012), post-tax fuel prices are used to proxy for carbon tax. Since fuel prices are available only at country level, we apply the following formula to exploit the yearly cluster-level variation in the dependent variable:[16]

$$FP_{i,t} = \sum w_{i,c} FP_{t,c}$$

where FP_t represents the tax-inclusive fuel price defined as the average of the diesel and gasoline price; and $w_{i,c}$ is a cluster-specific weight which captures the importance of country c in both green and non-green clusters. Therefore, for each cluster, we define the weight of country c according to the origin of the assignees and the number of their patents in the cluster. The higher the percentage of patents filed in country c, the greater w_c. To avoid potential sources of endogeneity deriving from the correlation between patents and fuel prices (Popp, 2002), we calculate w as a time-invariant weight using data over the whole period 1982-2008. Also, since the production of inventions in the automotive industry is concentrated in three main geographical areas, c corresponds to EU, JP and US[17] (Figure 5 shows the trends in the average of the fuel and diesel total prices in EU, JP and US). Therefore, FP_c includes the Japanese and American fuel prices and the average fuel price for the European countries.

4.2.2 Technological proximity

Substitution between the two fields may be driven by the characteristics of the technology space. The variable $PROX$ is included to test the effect of relatedness between technological fields. In the search for new knowledge, firms use routinised behaviour to search the closest knowledge fields to reduce the uncertainty of the process (Boschma, 2005).

[16]Aghion et al. (2012) exploit firm-level variation using the firm share of patents filed in country c.
[17] Different country level fuel prices were tested to make our results more robust.

Nelson and Winter (1982) emphasise that what emerges when firms search for new knowledge is often uncertain and unexpected. Therefore, the technological opportunities identified within clusters at a lower cognitive distance may induce firms to consider those technological fields as potential sources of knowledge with lower uncertainty. Hence, competition between two clusters might be explained by their cognitive proximity in the sense that a closer knowledge base may provide opportunities for further improvements to the technological field under investigation.

The literature suggests different ways to measure cognitive distance. Using a matrix and tracing R&D expenditure from the industry of origin to the industry of use for the resulting products and services, Scherer (1982) assumes that two industries can be considered close if a high share of the R&D performed in one industry is exploited by the other. In contrast to a user-producer-oriented methodology, the co-occurrence of classification codes within a patent document is employed to identify the relationship between the knowledge bases in different patent fields. The assumption is that the co-occurrence of patent classes measures the strength of the knowledge link and the spillovers between technological areas. Jaffe (1986) calculates the distribution of patents over 49 technological fields, for a sample of US firms, and measures the correlation (angular separation) between the research efforts performed in each innovative area to obtain the similarity between firms' R&D activities through a cosine index.

Following Jaffe (1986), we calculate the distance between cluster centroids in the technology space defined above (Figure 3(b)) and employ it as a proxy for knowledge relatedness between technological fields. In order to exploit cluster-pairwise variation of our dependent variable, we calculate the proximity ($PROX$) between technological efforts as follows:

$$PROX_{i,t} = \frac{PAT_{ng,t}}{DIST_i}$$

where technological proximity between each pair of clusters (i) is equal to the number of patents in the non-green cluster ($PAT_{ng,t}$) divided by the distance between the centroids of the two clusters ($DIST_i$) in the Euclidean space. Let $j = (j_1, j_2)$ be the coordinate of the non-green cluster centroid, and $k = (k_1, k_2)$ the coordinates of the green cluster centroid, then the distance between j and k is calculated as follow:

$$DIST_i = \sqrt{(k_1 - j_1)^2 + (k_2 - j_2)^2}$$

This formula allows us to weight the knowledge included in non-green clusters by its similarity to the knowledge in the green cluster. Thus, firms might deter inventive activities at relative small (large) cognitive distances, which implies that the search process is carried out among similar (different) technological fields. For example, inventive efforts in new promising environmental technology fields may reduce patenting in technology fields that are related to the internal combustion engine or conventional vehicle components.

4.2.3 Other variables

We also include variables to capture the linkage between cluster knowledge bases. In order to hold constant other aspects which might influence the propensity to substitute efforts in two technology fields, we control for the number of citations among technological fields (CIT) and the number of firms that file patents in each pair of clusters (NoF). The former aims to detect the technical relationship between technological domains from a vertical perspective since, patents include citations to earlier patents, which subsequent patent applications build on (OECD, 2009). This is a good indicator of the previous knowledge used by inventors for their inventions (Popp, 2002). The citation count for each pair of clusters is used to build the CIT variable. Note that this variable differs from

$PROX$ in the same way as knowledge similarity differs from knowledge flow. While cognitive distance detects proximity among clusters (within the whole dataset), citations identify the extent to which past knowledge embodied in a technology cluster is exploited by others. Hence, the CIT variable is closer to the concept of vertical complementarity since the generation of new knowledge is conditional on identification and integration of different complementary 'modules' in which recombination plays a pivotal role (Antonelli, 2003). Along these lines citations track the recombination of pieces of knowledge acquired in the past with recently acquired knowledge.

We focus also on the current relationship among the technological knowledge in different clusters, based on the number of firms that patent across clusters. We assume that if firms exploit more than one invention in a different technological field, this can be interpreted as a relationship between the knowledge bases in those clusters. We draw here on the concept of knowledge compositeness to interpret knowledge inter-dependence between technological fields. Knowledge compositeness is defined as the 'variety of units of technological knowledge that are necessary and complementary in the production of a new product or process, as well as of a new unit of knowledge' (Antonelli and Calderini, 2008, p. 24). From the perspective of the automotive industry, the importance of knowledge compositeness highlights that the challenge posed by the technological and scientific advances in the industry no longer reside in a single technological field (Antonelli and Calderini, 2008).

Finally, we include the stock of patents in environmental and non-environmental technology fields (PS). Aghion et al. (2012) highlight that prior knowledge affects the propensity to innovate in green and non-green technologies due to the presence of lock-in effects. Following Cockburn and Griliches (1988), Peri (2005) and Aghion et al. (2012), we calculate the stock of patents in each cluster using the perpetual inventory method:

$$PS_{z,t} = PAT_{z,t} + (1 - \delta)PS_{z,t-1}$$

where PS is the patent stock in each cluster z[18] and PAT its patent count in each year. Following the related literature, we set the depreciation of R&D capital (δ) at 20%.

4.3 The empirical model

The empirical model that we estimate to investigate what affects the competition between inventive efforts in different technological fields is:

$$CO_{i,t} = \beta_1 FP_{i,t-1} + \beta_2 PROX_{i,t-1} + \beta_3 CIT_{i,t-1} + \beta_4 NoF_{i,t-1} + \beta_5 PS_{g,t-1} + \beta_5 PS_{ng,t-1} + \alpha_i + \gamma_{i,t} + \epsilon_{i,t}$$

where FP is the post-tax fuel prices. We check also whether technological relatedness provides an incentive to shift from ng to g technological efforts, by including the $PROX$ variable. This variable captures the technological similarity between each pair of clusters, weighting the number of non-green patents by the pairwise distance between green and non-green clusters calculated in the output map shown in Figure 3. We control also for those factors which might influence the propensity to reduce inventive effort in one technological field in favour of another, that is, the number of citations (CIT), the number of firms that patent in each pair of clusters (NoF) and the stock of past knowledge in each cluster (PS). Finally, fixed effects α_i captures unobservable cluster-pairwise-specific time invariant heterogeneity and $\gamma_{i,t}$ is the cluster-pairwise-specific time trend which accounts for unobservable factors associated with each pair of clusters.

[18] This variable takes account of the patent stock in each technology field. The resulting variables employed in the analysis are PS_g and PS_ng which identify the patent stock in green and non-green clusters for each observation. When we test for the presence of a potential crowding out effect between green technological efforts, we use PS_g1 and PS_g2 to identify two patent stocks - one for each green cluster.

5 Results

In this section, we present and explain the findings for both hypotheses tested (green vs non-green and green vs green inventive activities) over the period 1982-2008 (27 years). Table 2 presents the descriptive statistics and Table 3 presents correlation matrices for both the models.

5.1 Green vs non-green patents

The results of the fixed effects linear model showing competition between green vs non-green inventive activities are shown in Table 4 column 2. The independent variables are lagged one year to account for the time to exploiting an invention,[19] which is common practice in the related literature (Aghion et al., 2012; Lee et al., 2011; Popp and Newell, 2012).

In analysing the results, we observe that an increase in tax-inclusive fuel prices enhances the likelihood that green inventions come at the expense of non-green ones. Since environmental regulation triggers automotive environmental inventive efforts (Aghion et al., 2012; Lee et al., 2012; Hascic et al., 2009), these results would suggest that fuel prices influence firms' incentives to reallocate patenting efforts from non-green to green technology fields.

This result can be interpreted as showing that post-tax fuel prices have an impact on the competition between the two technological fields and contribute to crowding out ICEV inventive activities in favour of alternative propulsion vehicle technologies. This suggests that tax-inclusive fuel prices would be effective at unlocking the automotive technological system from dependence on conventional vehicle innovations. Higher fuel prices encourage firms to conduct environmental invention activity and discourage development of dirty inventions (Aghion et al., 2012). Also, a fuel price increase reduces both non-environmental patent efforts and the social benefits that arise from eco-innovation.

[19] Note that we collected patents using the earliest priority year in the patent family, which indicates the first application for the patent at any patent office. This date is the closest to the end of the invention process which makes it unnecessary to include additional lags to account for patent office administrative procedures (18 months on average for patent award).

The results for the proximity variable have other implications. The coefficient indicates that the greater the proximity between technological clusters (i.e., the closer the clusters in the technology space), the smaller the shift required from non-environmental to environmental inventive activities. Thus, firms in technology clusters will tend to reduce efforts too distant (i.e., in terms of CPC classes) from green activities. Figure 3 shows that more distant technological clusters (the upper part of the map) with respect to green clusters, are not related directly to internal combustion engines. This highlights that, holding the other variables constant, firms' patenting strategies are directed towards increasing environmental innovation efforts at the expense of non-green inventions such as tyres, suspension, etc., or, alternatively, that this effect is higher for those clusters located at a greater distance in the technology space. That is, the efforts made to develop alternatively powered engines (such as hybrid, electric or fuel cell), seem to compete with development of conventional car components rather than with fossil fuel engine technologies.

5.2 Green vs green patents

We investigated the potential effect of fuel prices on the competition among green technologies. This is fundamental for testing whether green inventions drive out invention efforts in other green technological fields due to the increase in fuel prices or other factors which influence technological competition. We account for the effect of each green technological cluster on the others in the green domain. Table 4 column 3 shows that the coefficient of fuel price is positive and significant. This suggests that tax inclusive fuel prices may redirect technological efforts towards other environmental domains, increasing the likelihood of potential lock-in to sub-optimal alternative technologies. Rather than inducing improvements in a particular technological field, tax-inclusive fuel prices should encourage firms to exploit a variety of technological trajectories (Frenken et al., 2004) since, at the current stage of technological development, it is difficult to assess whether an alternative technology is superior. For example, although fuel cell vehicles are considered the most promising technology compared to hybrid and electric cars, there are some important bottlenecks which require resolution and, therefore, the risk remains of lock-in a technology which ultimately might be sub-optimal (Frenken et al., 2004).

In this regard, from a descriptive perspective we tested the relationships among patents in the green technology domain. This provides an explanation for the findings in the last column of Table 4 that fuel prices seem to redirect technological efforts among green technological fields. In Section 3.3, we showed that the main trajectories characterising green patenting efforts are related to technologies to improve the environmental performance of conventional vehicles, HVs, EVs and FCVs. Therefore, we group green clusters into these four categories. In line with Popp and Newell (2012), Table 5 shows a correlation between the percentage of patents per year in each category over three time spans.[20] The highest negative correlation is between conventional vehicle green technologies and other vehicle propulsion technologies, in all periods. This suggests that competition among green patenting activities relates mainly to these two broad categories of technological effort that is, the greening of conventional vehicles vs EVs, HVs and FCVs. Thus, even within the environmental technology domain, there is competition between conventional and alternative vehicle designs. However, correlation matrices do not enable us to identify the direction of this effect, which should be explored in future research.

In Table 4 column 3The proximity variable is positive and statistically insignificant, meaning that technological relatedness does not influence a shift between green technology efforts. Indeed, firms respond to the technological opportunities from continuous technological advances, highlighting the absence of a dominant technology among the alternatives to the fossil fuel engine. Since environmental patenting efforts occur in a variety of technological fields, the dynamic changes in these technological trajectories induce firms to invest in a portfolio of environmentally-friendly technologies with higher technological opportunities. However, in this case, the similarity between technological activities does not influence the shift from one technological field to another.

[20] Cut points are chosen to create three time spans that include the same number of years and allow us to investigate the correlation between green patenting activities over different periods. The result holds for different time spans.

5.3 Robustness analysis

We conducted some robustness checks to assess the reliability of the model using different variables. Table 6 shows the results employing 3, 4, and 5 year patent count moving averages as the dependent variable. The main results are based on a 4-year moving average, but the signs of the coefficients and their significance are almost unchanged - at least for the main independent variables - if we use 3 or 5 year patent count moving averages in the specifications.

We also ran the model using a different variable to capture the level of fuel prices (Table 7). The results in Table 4 are based on tax-inclusive fuel prices in three main areas (i.e. EU, JP and US); Table 7 presents the model results using the set of countries in each cluster[21] (FP_all). This variable is obtained by calculating the share of patents from each country of origin, in each cluster, multiplied by the tax-inclusive fuel price in each country area. Again, the signs and significance of the coefficients are almost the same with either variable.

Finally, Table 8 presents the results for fuel taxes ($FTax$) rather than tax-inclusive fuel prices. However, due to fuel tax data availability, the period of analysis is reduced (1986-2008). Also in this case we obtain similar results using tax-inclusive fuel prices and fuel taxes. This finding provides an insight into the effectiveness of fuel tax for fostering competition among alternatives.

[21] AT, DE, FR, IT, JP, KR, SE, UK, US

6 Conclusions

In this paper we analysed the dynamics of the invention activities pursued by large automotive firms with a specific focus on the role of prices in influencing competition between conventional and low-emissions vehicle technologies. This study contributes to the literature by explicitly examining whether fuel prices (used as a proxy for carbon tax) and technological proximity foster the substitution of non-green patents by green ones. We contribute also by analysing whether substitution between technological efforts related to alternative vehicles is favoured by these factors and provide evidence of their impact on competition among low-emissions vehicles.

Our findings suggest that both tax-inclusive fuel prices and fuel taxes - praying for carbon tax - induce a shift from non-environmental inventive efforts towards activities related to the development of alternative vehicles. The study provides insights into the crowding out effect of fuel prices, which favours substitution rather than complementarity among inventive efforts. We highlighted the effectiveness of this mechanism for delinking the automotive technological system from fossil fuel dependent technologies.

In addition to fuel prices, other factors affect competition with technological similarity between green and non-green clusters playing a pivotal role. We show that technological proximity has a negative impact on the shift from non-green to green inventions, and that environmental technologies related to hybrid, electric and fuel cell vehicles compete with conventional vehicle components which are located at a greater distance in the technology space. Thus, substitution seems mainly to affect technologies not directly related to the powertrain system. We tested the hypothesis that environmental policies influence competition among alternative technological efforts. The results would suggest that tax-inclusive fuel prices affect competition between environmental technological domains. This could increase the risk of lock-in to suboptimal substituting vehicle technologies because, at the current stage of development, the technology community cannot identify a best alternative to the internal combustion engine. We observed that this effect can affect green invention activity and environmental technologies related to fossil fuel vehicles. Even within the environmental technological domain there is competition between low emissions vehicle technologies and the 'greening' of conventional vehicle designs. However, further investigation is needed to assess the direction of this potential shift. In other words, if alternative vehicle inventions crowd out technological efforts that reduce the environmental impact of conventional cars, the likelihood of unlocking the automotive industry from its dependence on fossil fuel will increase. Steering invention efforts away from long run (development of alternative powertrain systems) to short run technological solutions (catalytic converters, improved efficiency of conventional engines, etc.) would reduce the capability of the automotive industry to escape lock in to the internal combustion engine .

References

Acemoglu, D., Aghion, P., Bursztyn, L., & Hemous, D. (2012). The Environment and Directed Technical Change. The American Economic Review, 102(1), 131-166.

Aghion, P., Dechezlepretre, A., Hemous, D., Martin, R., and Reenen, J. V. (2012). Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry. Working Paper 18596, National Bureau of Economic Research.

Allcott, H., & Wozny, N. (2014). Gasoline prices, fuel economy, and the energy paradox. Review of Economics and Statistics, 96(5), 779-795.

Antonelli, C. (2003). Knowledge complementarity and fungeability: implications for regional strategy. Regional Studies, 37(6-7), 595-606.

Antonelli, C., & Calderini, M. (2008). The governance of knowledge compositeness and technological performance: the case of the automotive industry in Europe. Economics of Innovation and New Technology, 17(1-2), 23-41.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. The economic journal, 116-131.

Barbieri, N. (2015). Investigating the impacts of technological position and European environmental regulation on green automotive patent activity. Ecological Economics, Volume 117, September 2015, Pages 140–152

Barbieri, N., Ghisetti, C., Gilli, M., Marin, N., Nicolli, F. (2016). A Survey of the Literature on Environmental Innovation Based on Main Path Analysis. Journal of Economic Surveys. Article first published online: DOI: 10.1111/joes.12149

Bergek, A., Berggren, C., & KITE Research Group. (2014). The impact of environmental policy instruments on innovation: A review of energy and automotive industry studies. Ecological Economics, 106, 112-123.

Boschma, R. (2005). Proximity and innovation: a critical assessment. Regional studies, 39(1), 61-74

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. Research Policy, 32(1), 69-87.

Bruckner, E., Ebeling, W., Montaño, M. J., & Scharnhorst, A. (1996). Nonlinear stochastic effects of substitution - an evolutionary approach. Journal of Evolutionary Economics, 6(1), 1-30.

Busse, M. R., Knittel, C. R., & Zettelmeyer, F. (2009). Pain at the pump: the differential effect of gasoline prices on new and used automobile markets (No. w15590). National Bureau of Economic Research.

Cecere, G., Corrocher, N., Gossart, C., & Ozman, M. (2014). Lock-in and path dependence: an evolutionary approach to eco-innovations. Journal of Evolutionary Economics, 24(5), 1037-1065.

Cockburn, I., & Griliches, Z. (1988). Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents. The American Economic Review, 419-423.

Cohen, W. M., & Klepper, S. (1996). A reprise of size and R & D. The Economic Journal, 925-951.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not) (No. w7552). National Bureau of Economic Research.

Costantini, V., Crespi, F., & Curci, Y. (2013). BioPat: An Investigation Tool for Analysis of Industry Evolution, Technological Paths and Policy Impact in the Biofuels Sector. In The Dynamics of Environmental and Economic Systems (pp. 203-226). Springer Netherlands.

Cowan, R., & Hultén, S. (1996). Escaping lock-in: the case of the electric vehicle. Technological forecasting and social change, 53(1), 61-79.

Crabb, J. M., & Johnson, D. K. (2007). Fueling the Innovative Process: Oil Prices and Induced Innovation in Automotive Energy-Efficient Technology. Technical report, Colorado College Department of Economics and Business, Colorado Springs, CO.

David, P. A. (1985). Clio and the Economics of QWERTY. The American economic review, 332-337.

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. Pattern Analysis and Machine Intelligence, IEEE Transactions on, (2), 224-227.

Dernis, H. and M. Khan (2004), "Triadic Patent Families Methodology", OECD Science, Technology and Industry Working Papers, 2004/02, OECD Publishing. http://dx.doi.org/10.1787/443844125004

Di Maria, C., & Van der Werf, E. (2008). Carbon leakage revisited: unilateral climate policy with directed technical change. Environmental and Resource Economics, 39(2), 55-74.

Dijk, M., & Yarime, M. (2010). The emergence of hybrid-electric cars: innovation path creation through co-evolution of supply and demand. Technological Forecasting and Social Change, 77(8), 1371-1390.

Faber, A., & Frenken, K. (2009). Models in evolutionary economics and environmental policy: Towards an evolutionary environmental economics. Technological Forecasting and Social Change, 76(4), 462-470.

Frenken, K., Hekkert, M., & Godfroij, P. (2004). R&D portfolios in environmentally friendly automotive propulsion: variety, competition and policy implications. Technological Forecasting and Social Change, 71(5), 485-507.

Frietsch, R., & Schmoch, U. (2006). Technological structures and performance as reflected by patent indicators. In National Systems of Innovation in Comparison (pp. 89-105). Springer Netherlands.

Gelijns, A. C., Zivin, J. G., & Nelson, R. R. (2001). Uncertainty and technological change in medicine. Journal of Health Politics, Policy and Law, 26(5), 913-924.

Gray, W. B., & Shadbegian, R. J. (1998). Environmental regulation, investment timing, and technology choice. The Journal of Industrial Economics, 46(2), 235-256.

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

Hascic I., de Vries F., Johnstone N. and Medhi N., (2009). Effects of environmental policy on the type of innovation: The case of automotive emission-control technologies. OECD Journal: Economic Studies. OECD Publishing, vol. 2009(1), pages 1-18.

Hottenrott, H., & Rexhäuser, S. (2013). Policy-induced environmental technology and inventive efforts: Is there a crowding out? (No. 128). DICE Discussion Paper.

Jaffe, A. B. (1986). Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. The American Economic Review, 76(5), 984-1001.

Jaffe, A. B., Newell, R. G., & Stavins, R. N. (2002). Environmental policy and technological change. Environmental and resource economics, 22(1-2), 41-70.

Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9), 1464-1480.

Kohonen, T. (1998). The self-organizing map. Neurocomputing, 21(1), 1-6.

Kohonen, T. (2001). Self-organizing maps (Vol. 30). Springer Science & Business Media.

Kohonen, T. (2013). Essentials of the self-organizing map. Neural Networks, 37, 52-65.

Kneller, R., & Manderson, E. (2012). Environmental regulations and innovation activity in UK manufacturing industries. Resource and Energy Economics, 34(2), 211-235.

Lanjouw, J. O., Pakes, A., & Putnam, J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. The Journal of Industrial Economics, 46(4), 405-432.

Laperche, B., Lefebvre, G., & Langlet, D. (2011). Innovation strategies of industrial groups in the global crisis: Rationalization and new paths. Technological forecasting and social change, 78(8), 1319-1331.

Lee, J., Veloso, F. M., & Hounshell, D. A. (2011). Linking induced technological change, and environmental regulation: Evidence from patenting in the US auto industry. Research Policy, 40(9), 1240-1252

Leydesdorff, L., Alkemade, F., Heimeriks, G., & Hoekstra, R. (2015). Patents as instruments for exploring innovation dynamics: geographic and technological perspectives on "photovoltaic cells". Scientometrics, 102(1), 629-651.

MacQueen J. B. (1967). Some methods for classification and analysis of multivariate observations.In Proceedings of 5th Berkeley symposium on mathematical statistics and probability, pp. 281–297.

Mansfield, E., Rapoport, J., Romeo, A., Wagner, S., & Beardsley, G. (1977). Social and private rates of return from industrial innovations*. The Quarterly Journal of Economics, 221-240.

Marin, G. (2014). Do Eco-Innovations harm productivity growth through crowding out? Results of an extended CDM model for Italy. Research Policy, 43(2), 301-317.

Martinelli, A., & Nomaler, Ö. (2014). Measuring knowledge persistence: a genetic approach to patent citation networks. Journal of Evolutionary Economics, 24(3), 623-652.

Nelson, R.R. and Winter, S.G. (1982). An evolutionary theory of economic change. Cambridge Mass: Harvard University Press.

Nesta, L., & Saviotti, P. P. (2005). Coherence of the knowledge base and the firm's innovative performance: evidence from the us pharmaceutical industry. The Journal of Industrial Economics, 53(1), 123-142.

Noailly, J., & 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.

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

Oltra, V., & Saint Jean, M. (2009). Variety of technological trajectories in low emission vehicles (LEVs): a patent data analysis. Journal of Cleaner Production, 17(2), 201-213.

Pakes, A. (1985). On Patents, R & D, and the Stock Market Rate of Return. The Journal of Political Economy, 390-409.

Peri, G. (2005). Determinants of knowledge flows and their effect on innovation. Review of Economics and Statistics, 87(2), 308-322.

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

Popp, D. (2005). Lessons from patents: using patents to measure technological change in environmental models. Ecological Economics, 54(2), 209-226.

Popp, D., Newell, R. G., & Jaffe, A. B. (2010). Energy, the Environment, and Technological Change. Handbook of the Economics of Innovation, 2, 873-937.

Popp, D., & Newell, R. (2012). Where does energy R&D come from? Examining crowding out from energy R&D. Energy Economics, 34(4), 980-991.

Porter, M. E., & Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. Journal of economic perspectives, 9, 97-118.

Rennings, K., Markewitz, P., & Vögele, S. (2013). How clean is clean? Incremental versus radical technological change in coal-fired power plants. Journal of Evolutionary Economics, 23(2), 331-355.

Scherer, F. M. (1982). Inter-industry technology flows in the United States. Research policy, 11(4), 227-245.

van den Bergh, J. C. (2013). Environmental and climate innovation: Limitations, policies and prices. Technological Forecasting and Social Change, 80(1), 11-23.

Thoma, G., Torrisi, S., Gambardella, A., Guellec, D., Hall, B. H., & Harhoff, D. (2010). Harmonizing and combining large datasets–an application to firm-level patent and accounting data (No. w15851). National Bureau of Economic Research.

Figures

Figure 1 – Green and non-green patent trends and percentage of green patents over total yearly patents

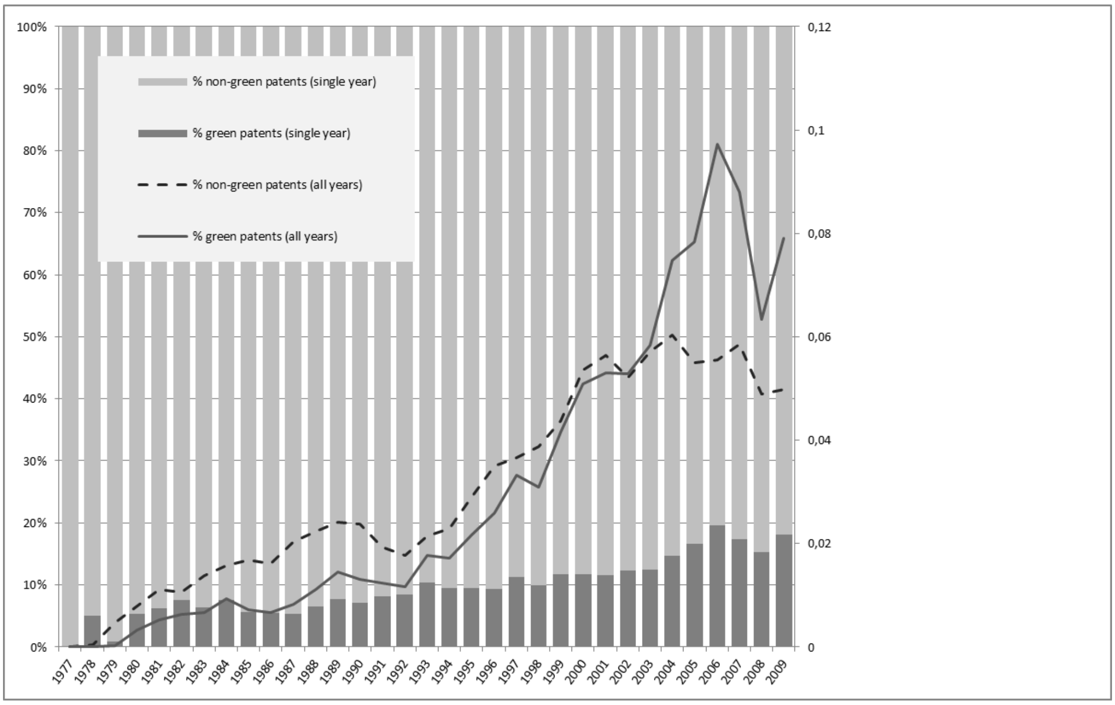


Figure 2 – Steps of SOM algorithm

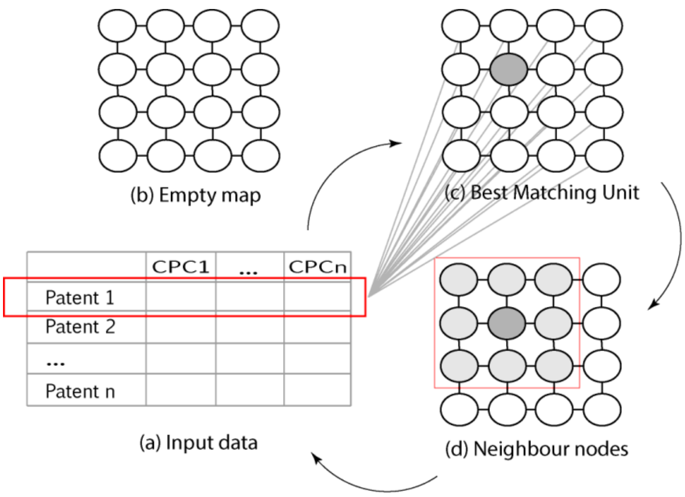
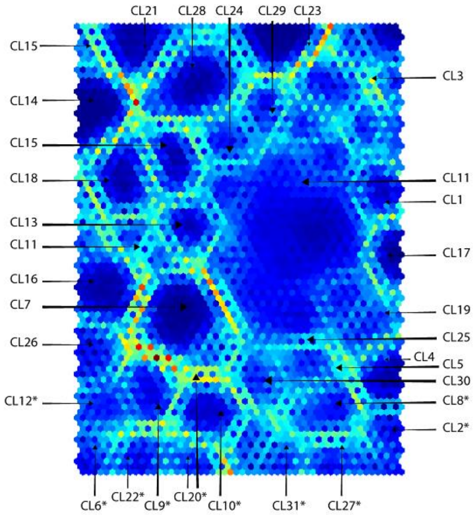


Figure 3 – Unified-distance matrix and clustering results

a)



b)

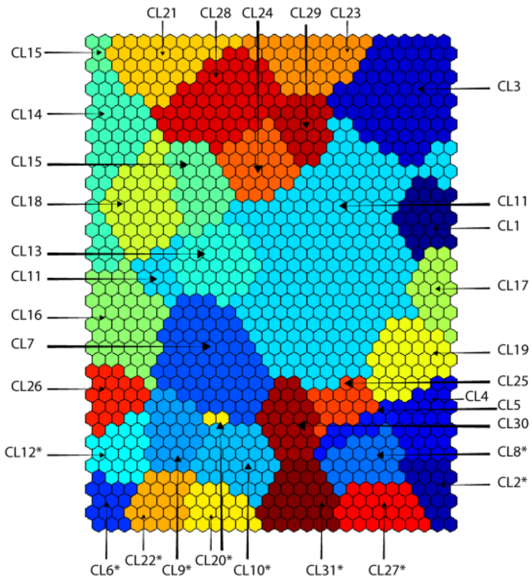


Figure 4 – Number of citations between patents in each cluster. Citations within each cluster are reported in the diagonal

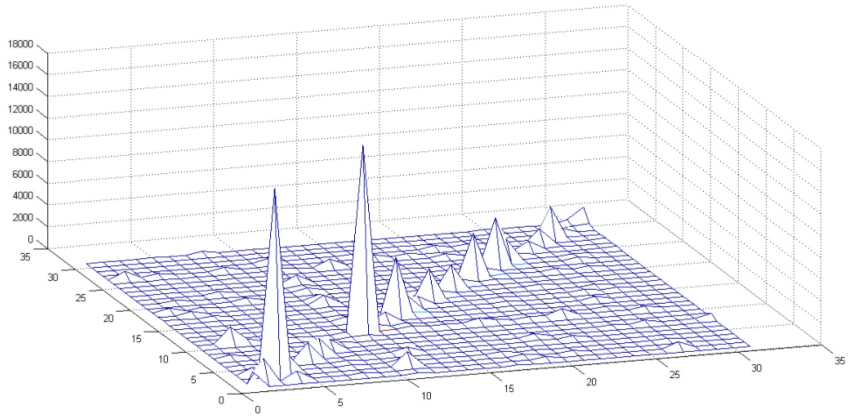
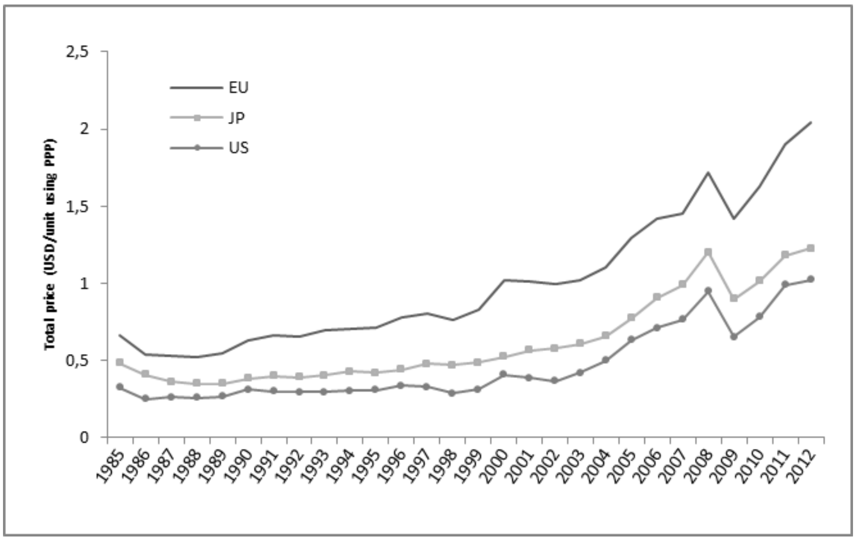


Figure 5 – Average post-tax fuel price between premium unleaded 95 RON and diesel in OECD countries, Japan and United States



Source: Own figure using data from IEA

Tables

Table 1 – Description of clusters

CL	Keywords	Number of patents	% of green patents
1	Bore, crank, pistons	760	0,00
2	**Ignition, catalyst, throttle**	**1455**	**100,00**
3	Tyre, rubber, pneumatics	5328	1,33
4	Injector, spark, crank	1640	0,00
5	NOx, SOx, particulate	98	0,00
6	**Battery, hybrid, regeneration**	**699**	**99,86**
7	Cell, cathode, anode	1598	0,31
8	**NOx, catalyst, purification**	**514**	**98,25**
9	**Gear, stator, transmission**	**174**	**100,00**
10	**Spark, battery, octane**	**347**	**100,00**
11	Wiper, door, antenna	18568	0,15
12	**Transmission, gear, hybrid**	**359**	**99,44**
13	Stator, pole, rotor	882	0,68
14	Transmission, pulley, hydraulic	2864	1,26
15	Caliper, friction, brake	1330	0,90
16	Pointer, drowsiness, menu	1091	0,00
17	Injector, nozzle, carburetor	1673	7,23
18	Rubber, etch, windscreen	1076	0,37
19	Camshaft, rocker, crankcase	1445	0,14
20	**Hydrogen, electrolyte, cell**	**525**	**100,00**
21	Brake, master, skid	2047	0,64
22	**Battery, charger, PLC**	**502**	**99,60**
23	Airbag, inflate, retractor	2841	0,32
24	Suspensions, strut, axle	1001	0,00
25	Muffler, catalyst, silencer	357	0,00
26	Cruise, yaw, headway	604	0,00
27	**Turbocharger, supercharger, swirl**	**1115**	**100,00**
28	Robot, crawler, roof	1758	0,23
29	Rubber, flywheel, diaphragm	571	0,00
30	Oxide, palladium, acid	328	0,00
31	**Catalyst, NOx, purification**	**820**	**100,00**
	Total	54370	11.97

Note: Clusters classified as green are highlighted in bold. The percentage of green patents over the total number of patents considered in this study is reported in the last row.

Table 2 – Descriptive statistics

Gr. vs. Non-Gr

Variable	Obs	Mean	Std. Dev.	Min	Max
CO	*5670*	*-2.233774*	*11.27849*	*-80.25*	*43*
FP (t-1)	*5670*	*1.037521*	*.4561879*	*.1190692*	*2.337916*
PROX (t-1)	*5670*	*71.1831*	*161.6976*	*0*	*2412.939*
CIT (t-1)	*5670*	*.9640212*	*4.023495*	*0*	*82*
F (t-1)	*5670*	*3.183774*	*2.92954*	*0*	*18*
PS ng (t-1)	*5670*	*287.1834*	*579.0677*	*0*	*4769.544*
PS g (t-1)	*5670*	*68.0297*	*83.43386*	*0*	*478.0059*

Gr. vs. Gr

Variable	Obs	Mean	Std. Dev.	Min	Max
CO	*2430*	*2.834568*	*3.325596*	*0*	*24*
FP (t-1)	*2430*	*1.00259*	*.4640165*	*.1162741*	*2.200845*
PROX (t-1)	*2430*	*38.94946*	*87.48263*	*0*	*1050.328*
CIT (t-1)	*2430*	*2.453498*	*7.494261*	*0*	*98*
F (t-1)	*2430*	*2.394239*	*2.485058*	*0*	*12*
PS g (t-1)	*2430*	*68.0297*	*83.44367*	*0*	*478.0059*

Table 3 – Correlation matrix

Gr. vs. Non-Gr

Variable	1	2	3	4	5	6
FP (t-1)	*1*					
PROX (t-1)	*0.2137*	*1*				
CIT (t-1)	*0.0662*	*0.2000*	*1*			
NoF (t-1)	*0.6039*	*0.3861*	*0.2257*	*1*		
PS ng (t-1)	*0.2808*	*0.9053*	*0.1512*	*0.4451*	*1*	
PS g (t-1)	*0.6853*	*0.1352*	*0.1465*	*0.7121*	*0.2074*	*1*

Gr. vs. Gr

Variable	1	2	3	4	5	6
FP (t-1)	*1*					
PROX (t-1)	*0.3274*	*1*				
CIT (t-1)	*0.1413*	*0.4954*	*1*			
NoF (t-1)	*0.6954*	*0.5741*	*0.3592*	*1*		
PS g1 (t-1)	*0.6758*	*0.5881*	*0.2118*	*0.7287*	*1*	
PS g2 (t-1)	*0.6758*	*0.3728*	*0.2094*	*0.7287*	*0.4801*	*1*

Table 4 – Main results of fixed-effects linear model

	(1) Gr vs. non-Gr	(2) Gr. vs. Gr.
FP (t-1)	3.1682***	0.6117**
	(1.0034)	(0.2492)
PROX (t-1)	-0.2224***	0.0008
	(0.0341)	(0.0032)
CIT (t-1)	-0.1062	-0.0266**
	(0.0736)	(0.0114)
NoF (t-1)	0.0189	-0.0010
	(0.1138)	(0.0637)
PS g (t-1)a	0.0924***	0.0162***
	(0.0045)	(0.0041)
PS ng (t-1)	-0.0091	0.0175***
	(0.0055)	(0.0035)
_cons	490.1122***	312.96***
	(166.9565)	(94.2007)
N	5670	2430
r2	0.6498	0.4938

The two columns refer to the hypotheses tested above. Gr vs. non-Gr tests the dynamics of research efforts between green and non-green, whereas Gr vs. Gr tests the hypothesis that green research efforts drive away other green research efforts. Dependent variable: CO using 4-years moving average.

[a] In the first column the patent stock is calculated on green and non-green clusters. In the second column, even if we maintained same variable names, the clusters are both green.

Robust clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 – Correlation between percentage of patents per year in each environmental inventive activity.

	Greening of conventional vehicles	Alternative vehicles		
		EV	HV	FCV
Correlation matrix 1982-1990				
Greening of conventional vehicles	1.00			
EV	-0.94 (0.00)	1.00		
HV	-0.84 (0.01)	0.59 (0.09)	1.00	
FCV	0.16 (0.68)	-0.08 (0.83)	-0.31 (0.42)	1.00
Correlation matrix 1991-1999				
Greening of conventional vehicles	1.00			
EV	-0.27 (0.47)	1.00		
HV	-0.47 (0.20)	-0.70 (0.03)	1.00	
FCV	-0.51 (0.16)	-0.32 (0.40)	0.54 (0.14)	1.00
Correlation matrix 2000-2008				
Greening of conventional vehicles	1.00			
EV	-0.67 (0.05)	1.00		
HV	-0.66 (0.05)	0.06 (0.88)	1.00	
FCV	-0.30 (0.44)	-0.21 (0.58)	0.07 (0.85)	1.00

Table 6 – Model results using different moving averages (3, 4, 5 years)

	Gr vs. non-Gr			Gr vs. Gr		
	3 Years MA	4 Years MA	5 Years MA	3 Years MA	4 Years MA	5 Years MA
FP (t-1)	2.2945**	3.1682***	4.4813***	0.9014**	0.6117**	-0.1426
	(0.9344)	(1.0034)	(1.0373)	(0.3506)	(0.2492)	(0.2194)
PROX (t-1)	-0.2438***	-0.2224***	-0.1914***	0.0000	0.0008	0.0064***
	(0.0370)	(0.0341)	(0.0293)	(0.0036)	(0.0032)	(0.0019)
CIT (t-1)	-0.1447	-0.1062	-0.0773	0.0018	-0.0266**	-0.0128
	(0.0928)	(0.0736)	(0.0627)	(0.0119)	(0.0114)	(0.0089)
NoF (t-1)	0.0583	0.0189	0.0328	-0.0640	-0.0010	0.0361
	(0.1405)	(0.1138)	(0.1017)	(0.0817)	(0.0637)	(0.0510)
PS g (t-1)	0.1175***	0.0924***	0.0681***	0.0138***	0.0162***	0.0067**
	(0.0049)	(0.0045)	(0.0040)	(0.0039)	(0.0041)	(0.0032)
PS ng (t-1)	-0.0273***	-0.0091	-0.0011	0.0137***	0.0175***	0.0114***
	(0.0062)	(0.0055)	(0.0053)	(0.0032)	(0.0035)	(0.0027)
_cons	379.5873**	490.1122***	343.6314**	-50.0743	312.9676***	192.1476***
	(188.5195)	(166.9565)	(153.7244)	(96.2058)	(94.2007)	(70.0258)
N	5670	5670	5670	2430	2430	2430
r2	0.6112	0.6498	0.6994	0.4098	0.4938	0.4391

*The two columns refer to the hypotheses tested above. Gr vs. non-Gr tests the dynamics of research efforts between green and non-green, whereas Gr vs. Gr tests the hypothesis that green research efforts drive away other green research efforts. Dependent variable: CO using 3, 4, 5 years moving averages. Robust clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

Table 7 – Model results using a different tax-inclusive fuel price variable

	Gr vs. Non Gr	Gr. vs. Gr
FP_all (t-1)	3.2838***	0.6573**
	(1.0009)	(0.2498)
PROX (t-1)	-0.2224***	0.0008
	(0.0341)	(0.0032)
CIT (t-1)	-0.1053	-0.0264**
	(0.0734)	(0.0114)
NoF (t-1)	0.0223	0.0008
	(0.1136)	(0.0636)
PS g (t-1)	0.0924***	0.0160***
	(0.0045)	(0.0041)
PS ng (t-1)	-0.0101*	0.0174***
	(0.0056)	(0.0035)
_cons	473.3723***	311.8109***
	(166.1476)	(93.6547)
N	5670	2430
r2	0.6499	0.4939

The two columns refer to the hypotheses tested above. Gr vs. non-Gr tests the dynamics of research efforts between green and non-green, whereas Gr vs. Gr tests the hypothesis that green research efforts drive away other green research efforts. Dependent variable: CO using a 4-years moving average. Robust clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8 – Main results of fixed-effects linear model using fuel taxes(1986-2008)

	(1) Gr vs. non-Gr		(1) GR vs. Gr	
FP (t-1)	3.0881**		2.7583***	
	(1.2072)		(0.7558)	
FTax (t-1)		11.9629**		12.0665***
		(5.5209)		(3.2177)
PROX (t-1)	-0.2309***	-0.2315***	0.0008	-0.0007
	(0.0314)	(0.0314)	(0.0035)	(0.0036)
CIT (t-1)	-0.1197*	-0.1267*	-0.0249**	-0.0279**
	(0.0716)	(0.0740)	(0.0110)	(0.0109)
NoF (t-1)	0.0389	-0.0025	0.0125	-0.0122
	(0.1124)	(0.1131)	(0.0711)	(0.0714)
PS g (t-1)	0.1071***	0.1084***	0.0065	0.0168***
	(0.0036)	(0.0033)	(0.0050)	(0.0042)
PS ng (t-1)	-0.0201**	-0.0030	0.0082*	0.0175***
	(0.0081)	(0.0058)	(0.0042)	(0.0036)
_cons	406.5658*	1294.4763***	-85.1920	724.8870***
	(244.7970)	(431.0234)	(130.9304)	(189.5299)
N	4830	4830	2070	2070
r2	0.6680	0.6674	0.4904	0.4875

The two columns refer to the hypotheses tested above. Gr vs. non-Gr tests the dynamics of research efforts between green and non-green, whereas Gr vs. Gr tests the hypothesis that green research efforts drive away other green research efforts. Dependent variable: CO using a 4-years moving average. Robust clustered standard errors in parentheses. $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$