Cleantech and policy framework in Europe: A machine learning approach

Annalisa Croce – Politecnico di Milano

Laura Toschi – Università degli Studi di Bologna

Elisa Ughetto – Politecnico di Torino

Sara Zanni – Università degli Studi di Bologna

Abstract

The pursuit of a sustainable and clean energy future has emerged as a paramount global imperative of the 21st century. Achieving this transition is a multifaceted and complex endeavor that requires a harmonious interplay of factors: effective policy frameworks, cleantech firms, and the transformative power of data science. By focusing on the European context, this paper advances the field in several directions. First, it explores the use of machine learning techniques to identify cleantech firms by analyzing their mission statements and addressing the weaknesses of the existing methods. Second, it collects a unique and comprehensive dataset of national-level policies addressing the different topics covered by the European Green Deal. Third, in a regression analysis at country level, it examines the interplay between the national regulatory framework and the birth and growth of the cleantech landscape, by distinguishing between innovators (firms which develop the cleantech) and ecosystem firms (which adopt the cleantech). Our results indicate that the introduction of policies favors by itself the birth of cleantech innovator companies and their growth in the country. An increasing number of policies has a regulatory effect in the cleantech ecosystem limiting the number of newborn companies while favoring their growth.

Keywords: Cleantech, Policy, Machine Learning, European Green Deal, Europe

1. Introduction

The pursuit of a sustainable and clean energy future has emerged as a paramount global imperative in the 21st century. In response to escalating concerns about climate change, resource depletion, and environmental degradation (IPCC, 2014), nations around the world are initiating various actions to support the implementation of a sustainable transition (Hiatt et al., 2015). The transition to clean energy sources, characterized by a reduced carbon footprint and increased reliance on renewable technologies, is a prerequisite for meeting these challenges. However, achieving this transition is not just a matter of technological innovation, but a multifaceted and complex endeavor that requires a harmonious interplay of factors. In this paper, we focus in particular on policy frameworks, cleantech firms, and the transformative power of data science.

The European Green Deal (EGD) was published in December 2019, in response to the declaration of a climate emergency. It is designed to define a broad strategy to address the challenges of climate change and sustainable development, limiting the trade-offs of environmental degradation, and their interlinkages (European Commission, 2019). Funded by a third of the €1.8 trillion investment in the NextGenerationEU Recovery Plan, the initiatives included in the package aim to reduce the continent's greenhouse gases emissions by 55% by 2030 and achieve carbon neutrality by 2050. By combining several elements (e.g. skills and competences, public and efficient transportation, healthy and affordable food, energy efficient buildings, clean energy, fresh air, clean water, healthy soil and biodiversity), the EGD aims to support the competitiveness of the EU economy, while ensuring both a decoupling between economic growth and resource depletion, and that no one is left behind, i.e. a just and sustainable transition (Mura et al., 2023).

National cleantech policies are an integral part of the EGD, as they help policymakers in EU member states to achieve the EGD's targets. In particular, the EGD highlights the importance of cleantech firms and their ability to develop innovative sustainability solutions. By providing the practical tools to implement the vision of the EGD, national cleantech policies should therefore be able to create a supportive ecosystem that promotes the creation and growth of cleantech firms by providing financial support, market incentives, and regulatory guidance, and by making it more attractive for entrepreneurs to enter the cleantech sector (Porter, 1991). At the same time, however, policymakers have a variety of approaches at their disposal (Grubb, 2004). An excessive number of complex and sometimes overlapping policies can potentially have a negative impact on the creation and operation of a cleantech ecosystem due to compliance burdens, uncertainty and regulatory risk, market fragmentation, regulatory barriers to entry, or lack of coordination, to name but a few of the possible drawbacks (Klapper et al., 2006). The relationship between cleantech policy and cleantech entrepreneurship is therefore intricate and requires a deep understanding of how policy decisions shape the cleantech landscape and influence its success and scalability. The strength and way in which EU countries have adopted regulatory frameworks to promote a clean transition has created a valuable setting for testing policy efficiency and effectiveness (Burer and Wustenhagen, 2009).

A proper assessment of the impact of policy interventions requires a precise and accurate identification of cleantech firms. Driven by innovation and sustainability, cleantech firms are at the forefront of developing and deploying cutting-edge technologies in the fields of renewable energy, energy efficiency, and environmental conservation. Indeed, they embody the spirit of disruption, challenging traditional energy paradigms and offering novel solutions that have the potential to revolutionize the energy landscape. As cleantech firms play a vital role in addressing environmental challenges, the ability to accurately identify them is critical for various stakeholders, including investors, policymakers, and sustainability advocates.

The use of Artificial Intelligence (AI) and data science provides a data-driven and automated approach to identifying a large number of cleantech firms with precision and accuracy. In particular, these approaches, with their advanced machine learning (ML) techniques, offer a powerful solution for identifying cleantech firms by automatically extracting relevant keywords and phrases from mission statements and identifying patterns and associations between specific words or phrases and cleantech activities. A company’s mission statement encapsulates its core values, goals, and purpose and, in the context of cleantech, includes key terms and phrases related to sustainability, environmental protection, and clean technologies. AI can therefore use this linguistic information to objectively classify firms as cleantech or non-cleantech on a large scale, overcoming the problems faced by previous studies.

Traditionally, the classification process has relied on manual assessment, which is time-consuming and subjective, or on predefined classifications based on industry labels (i.e., NACE or SIC codes), which have proven inefficient due to their inability to capture the heterogeneous nature of the cleantech sector, or on taxonomies (i.e., EU taxonomy), which due to their structural rigidity are unable to adapt to the dynamics of the industry (Christensen & Hain, 2017; Criscuolo & Menon, 2015; Cumming et al., 2016). Moreover, the terminology used to define cleantech firms is often ambiguous due to the lack of a universally accepted definition. This can lead researchers to neglect important entities in their analyses, resulting in a poor understanding of the true role and impact of policy frameworks for sustainability transitions. Finally, in the pursuit of profitability and market competitiveness, some firms engage in the common practice of "greenwashing", by classifying themselves as cleantech to give the false impression that they are more environmentally friendly and sustainable than they actually are (Ramus and Montiel, 2005; Walker and Wan, 2012). This practice has important consequences, as it can undermine the credibility of genuine cleantech firms and the broader sustainability movement. The use of AI-based techniques overcome all these drawbacks.

This academic article embarks on a comprehensive exploration of the complex interplay between clean energy transition, policy, cleantech firms, and data science. Through a multidisciplinary lens, we seek to unravel the synergies, challenges, and opportunities that emerge at this intersection and shed light on how these elements converge to shape the future of clean energy adoption and sustainability. In particular, our work contributes to the extant literature in three main ways. First, it explores the use of ML techniques to identify cleantech firms by analyzing their mission statements and addressing the weaknesses of the existing methods. Second, it collects a unique and comprehensive dataset of national-level policies addressing the different topics covered by the EGD. Third, it investigates the interplay between the national regulatory framework and the growth of the national cleantech ecosystem.

In the following, we first present a comprehensive overview of the main studies focused on analyzing the impact of policy regulations (Section 2). Then, we present the AI-based approach for identifying cleantech firms and the procedure used to identify our sample of cleantech and a control group of non-cleantech matched companies (Section 3) and the procedure for generating a comprehensive list of policies (Section 4). We then present the results (Section 5) and we conclude with policy implications and directions for future research (Section 6).

2. Background literature

Governments aim to increase the public benefits of sustainability by regulating firms to adopt sustainable practices (such as reducing pollution). As such, the institutional environment plays a crucial role in defining the boundaries of entrepreneurial opportunities and new venture creation. While it is well known that government regulation of firms' environmental responsibilities can be effective in limiting the detrimental effects of economic activity on the natural environment, a debate on the impact of environmental regulation on firm performance and new venture creation is still under scrutiny in the academic community (see, among others, Bi et al., 2014, Ford et al., 2014, Majumdar and Marcus, 2001, Porter and van der Linde, 1995a, Rubashkina et al., 2015). Indeed, the impact of the institutional environment on emerging industries can vary considerably. While it can encourage the creation of new firms through market design incentives, it can also discourage the creation of new firms through overly prescriptive and restrictive regulations (Bruton et al., 2010).

Two different perspectives have animated the debate on the impact of regulation on business creation and performance.

2.1. Regulation as a barrier to entrepreneurship

The first perspective is based on the common perception that regulation imposes a burden or constraint on firms, with the strongest effects on new firms compared to more established ones. According to this perspective, which is more widely accepted by institutional theorists, an institutional environment with too many rules and procedures can hinder the creation of new firms (Klapper et al., 2006).

Studies of compliance costs, for example, argue that regulation increases administrative and psychological costs for businesses, discourages investment, innovation and growth, and hinders the creation of new firms (Blackman et al., 2010; Chittenden et al., 2005; Gray and Shadbegian, 2003). Regulations are seen as harmful to business, despite being socially desirable, with small businesses suffering disproportionately (Kitching, 2006). In this area, entrepreneurs are at a competitive disadvantage to incumbents, which use their political power to steer policymakers towards a policy environment favourable to them rather than new entrants (Pacheco et al., 2010). This situation is referred to as the 'green prison', meaning that entrepreneurs are constrained in their ability to innovate and compete due to regulations that favour incumbents. In a recent study based on interviews with environmental entrepreneurs in the UK, France and Germany, Ball and Kittler (2019) find that the introduction of environmental policies and support mechanisms for low-carbon investment (such as market taxes and emissions trading schemes) were more effective in supporting large incumbents and utilities than in encouraging the creation of new businesses.

However, a perspective that considers regulation only as a constraint provides a limited understanding of the benefits that may accrue to businesses from the creation of market opportunities and improved efficiency and competitiveness. Regulations that legitimise certain business models can make it easier for entrepreneurs to start and grow their businesses. In addition, regulations that increase the disruptive potential of incumbent structures can create opportunities for new businesses to emerge and compete (Christensen et al., 2018). A number of studies have taken a more nuanced approach to regulation, recognising that regulation can also benefit firms.

2.2. Regulation as a facilitator of entrepreneurship

The concept that environmental regulations, rather than uniformly penalising all firms, may provide an opportunity for some firms to become more competitive and innovative and thereby improve their financial performance was introduced by Porter in 1991. The basic idea behind Porter's arguments is that environmental regulations (especially stricter ones) can induce efficiencies and encourage innovations that help to improve firms' economic competitiveness and business performance (Porter and van der Linde, 1995). Firms that are able to comply with regulatory requirements that are properly designed and sufficiently flexible can trigger the discovery and adoption of cleaner technologies, making production processes and products more efficient. This improved efficiency is responsible for cost savings sufficient to overcompensate both the compliance costs directly attributable to new regulations and the higher innovation costs (Ambec and Barla, 2006; Porter and van der Linde, 1995; Tello and Yoon, 2008). In the same vein, Johnstone et al. (2010) suggest that the stringency of environmental policies can provide incentives for innovation, while stable norms and standards can reduce uncertainty in investment decisions.

2.3. The effect of environmental regulation on entrepreneurship: a mixed scenario

Since the formulation of the Porter hypothesis (or win-win scenario), a significant number of papers have broadly tested this claim, but the results have been mixed (Ambec et al., 2013; Lanoie et al., 2011). In terms of firm performance, some studies have found that environmental regulation leads to lower financial returns (Filbeck and Gorman, 2004), while others have found a positive impact (Eiadat et al., 2008) or an insignificant relationship (Triebswetter and

Hitchens, 2005). Similarly, when the relationship between regulation and innovation has been examined, the results have been inconclusive (Jaffe and Palmer, 1997; Sanchez and McKinley, 1998; Shao et al., 2020), with the majority of studies finding a positive effect (Brunnermeier and Cohen, 2003; Horbach, 2008; Martinez-Zarzoso et al., 2019; Ramanathan et al., 2017). For example, Martinez-Zarzoso et al. (2019) show that, in the long run, stricter environmental policies lead to either an increase in innovative activity (in the form of R&D or patents) or an improvement in economic efficiency (as reflected in total factor productivity). Rennings and Rammer (2011) analyse the impact of regulation-driven environmental innovation on innovation success and firm performance. They find that regulation-driven environmental innovation generates innovation success similar to other types of innovation and has a positive impact on firm performance. More recently, a few studies have examined the relationship between regulation, innovation and performance simultaneously (Black et al., 2010; Lopez-Gamero et al., 2010; Hu et al., 2017; Montabon et al., 2007; Ramanathan et al., 2010; Triebswetter and Wackerbauer, 2008). Black et al. (2010) find that innovation positively moderates the relationship between regulation and economic performance in the case of more flexible regulation, but not in the case of less flexible regulation. Ramanathan et al. (2017) examine the relationship between environmental regulation, firm innovation and private sustainability benefits using nine case studies of UK and Chinese firms. The authors show that firms that take a more dynamic and proactive approach to responding to environmental regulations and managing their environmental performance are, on average, better able to reap sustainability benefits.

Recent studies have highlighted the importance of the design of regulations, the sector of reference and the ability of firms to implement regulations (Criscuolo and Menon, 2015, Sunny and Shu, 2017, and York and Lenox, 2014). For example, Kirkpatrick and Parker (2015) provide a critical literature review of the theory and quantitative evidence on the impact of regulatory policy. They argue that the impact of regulation on firm performance is complex and depends on a number of factors, including the type of regulation, the industry in which the firm operates, and the level of competition in the market.

In conclusion, the impact of policy and regulation on the emergence of new firms and their performance is complex and depends on a variety of factors. While some studies suggest that regulation can have a positive impact on new firm creation and firm performance, others argue that its impact is dynamic and depends on the specific context in which it is applied. Overall, the results of these studies suggest that policy makers should carefully consider the potential impact of policies and regulations on firms before designing them.

3. Method and data

3.1. A machine learning approach to identify cleantech firms in Europe

Given the shortcomings of existing methods for classifying cleantech firms (such as the NACE industrial classification and the EU taxonomy of sustainable investments), we developed an original and fully replicable methodology, based on supervised ML algorithms, to identify cleantech firms in Europe.

We started with the full sample of firms available in Orbis, a dataset maintained by Bureau Van Dijk that contains financial information on over 40 million firms worldwide and is widely used due to its company coverage, availability of financial data and potential for data harmonisation. We selected all European firms that had at least one financial year of accounting data and for which an extended business description was available. A total of 537,129 firms were identified.

By focusing on the business description, it is possible to determine the cleantech orientation of a firm more accurately than by relying on standard industry classifications. However, manual classification of firms based on their business description may not be feasible in the presence of very large databases. Therefore, we applied supervised ML algorithms to the extended business description of our sample to disentangle between cleantech and non-cleantech firms.

Broadly speaking, the process involved the following steps. In the first step, we manually analysed a relatively small set of firms identified as cleantech, which formed our training set. The training dataset was tuned by cross-validation, with the aim of allowing the machine to learn to identify cleantech firms based on what was written in the company's business description. To do this, a set of features (or predictors) were inferred from the text of the company business descriptions using text mining techniques, which allowed us to predict the classification of non-manually classified firms and filter out all firms that were definitely not cleantech. The second step was to apply a computerised filter to each firm identified as cleantech with the ML in order to eliminate false positives.

The supervised ML involved a number of phases, which are described in detail below. We first used text classification, a process that automatically assigns documents to one or more predefined categories based on their content (Yang and Liu, 1999). The method involves applying a supervised ML technique to a set of labelled documents (training set) to derive a decision function, which is then evaluated on another set of labelled documents (test set) and finally used to predict the category of new texts for which the classification is unknown. From the initial sample of 537,129 firms, we randomly selected a small sample of 8,501 firms to create the training and test sets. Two independent researchers looked at the company descriptions and classified the documents as cleantech or non-cleantech. Researchers were provided with a general definition of cleantech, which refers to technologies that aim to generate a positive environmental impact (e.g. in terms of reducing the consumption of non-renewable resources or the amount of waste generated (Pernick and Wilder, 2007)). Potential inconsistencies in the classification of a firm were discussed within the research team until agreement was reached. The documents were split to create the training and test sets according to a 70/30 rule (5,951 training and 2,550 test). We then generated a list of features using extraction techniques derived from text mining (Feinerer et al., 2008). This process involved text pre-processing, where a vector of documents is transformed into a corpus and later into a document term matrix (DTM). The algorithm is called "bag-of-words". We also used various techniques to clean up the text. We converted uppercase letters to lowercase, removed all unnecessary punctuation, numbers and symbols (e.g. @,°,# [,§ etc.), converted acronyms to regular expressions, e.g. "IT" to "Italy", checked spelling, replaced contractions, e.g. "I'm" to "I am". We also removed stop words, which are of little use for prediction.

After pre-processing the text, we converted words to their simplest form (text normalisation), reducing words to their common root. We developed a 'lemmatisation' process to analyse the different inflected forms of a word as a single lemma. The last step was to build a Document Term Matrix (DTM), with each document as a row and each n-gram (or term) as a column, and to calculate the frequencies (total number of times each n-gram appears in all documents), with the n-grams as the names of the vector. Two different tokeniser functions were created to construct the DTM for 1-gram and 2-gram. At the end of the process, we extrapolated 251 1-gram and 181 2-grams for a total of 432 features that were used to train the predictive model of being cleantech.

By allowing the machine to learn the mapping between the generated features and the 'true' cleantech label using the training dataset, it is possible to predict the classification of unlabelled firms and provide an appropriate predicted cleantech label. We used a number of popular ML methods (i.e., Naïve Bayes; Random Forest; Gradient Boosting Machines; Neural Network) and selected the classifier (GBM) whose prediction errors for non-manually classified firms were the smallest (best test accuracy, with a prediction accuracy greater than 90%). This accuracy was also robust to the presence of unbalanced data, due to the smaller number of cleantech firms in the training dataset than non-cleantech firms. For a detailed analysis of the performance of the cleantech classification prediction and the methods used to deal with the imbalance of the target variables (e.g., class weighting and random over-sampling examples (ROSE)), see Ambrois et al. (2023). The ML-based classification resulted in a sample of 74,047 firms (from an initial sample of 537,129 firms).

3.2. Computer-aided filters and manual classification

Secondly, we introduced a set of computational filters to reduce the number of false positives using functions embedded in the Stata software. We first analysed the cleantech literature to identify appropriate clean technology keywords, then validated them by looking at the company descriptions and searching for new additional keywords. Again, as in the ML step, we removed all punctuation, reduced the text to lower case, and examined the root of the word where necessary. After this additional process, our sample was reduced to 25,044 firms broadly defined as 'cleantech'.

As a third and final step we performed a manual classification to account for the complexity of the supply chain structure of the cleantech ecosystem. More specifically, this process aimed to classify each cleantech firm into ‘cleantech ecosystem’ (i.e., firms that adopt clean technologies, sell services based on clean technologies, or provide inputs for the development of clean technologies) and 'cleantech innovators’ (i.e., firms that are committed to develop clean technologies). To this end, two research assistants manually analysed the firms’ business descriptions to ensure maximum accuracy.

We thus identified ‘cleantech innovators’ as firms that create (and eventually use) the clean technology as their core business and are at the centre of the supply chain.

For the ‘cleantech ecosystem’ sample, we further distinguished such firms into “experimenters” and “manufacturers”, which support the realisation of the technology and “distributors”, “integrators”, and “operators”, which make the technology available to the market.[1]

Although each firm could fit into more than one definition, we decided to assign a unique class according to the firm’s primary activity.

Of the 25,044 cleantech firms identified, we classified 3,201 firms as cleantech innovators, focusing on the development of clean technologies, and the remaining 21,843 as cleantech ecosystem, referring to firms that adopt clean technologies, sell services based on clean technologies, or provide inputs for the development of clean technologies. [2]

3.3. Cleantech sample description

Table 1 reports the distribution of sample firms according to the segmentation described in Section 3.2. 13.30% of the firms are cleantech innovators. Within the cleantech ecosystem group, firms are equally distributed among integrators (27.2%), operators (22.8%) and manufacturers (22.3%). The two remaining groups (i.e., distributors and experimenters) account for 13.9% and 0.4% respectively.

[Insert Table 1 here]

More than half of the cleantech firms (51.18%) are located in only three countries: Germany (18.43%), Italy (17.16%), and France (14.09%), with the remaining firms distributed among

---

[1] Experimenters are firms involved in carrying out experimental tasks that can lead to discoveries and advances in the science of the cleantech supply chain (both private and public); Manufacturers are firms involved in the cleantech supply chain that provide ancillary services related to the actual innovation; in other words, they deal with the manufacturing, fabrication, and production of necessary and auxiliary components or raw materials for the clean technology; Distributors are firms that only distribute or are involved in the commercial provision of specific cleantech products or technologies. Their primary role is to bring clean technologies to market; Integrators are firms involved in the cleantech supply chain, that provide ancillary services related to the actual innovation (i.e., engineering, installation, procurement, design, conception, and planning). Their prominent role is to make the clean technology available to the users. Operators are firms involved in the cleantech supply chain that deal with the construction, implementation, and maintenance of facilities where the clean technology is used; in other words, they are ancillary services to the actual innovation. In addition, adopters that use technology as a primary tool for the achievement of their output (e.g. energy production) are also operators.

[2] The remaining 976 firms were not included in any category because it was not possible to assign them to a specific category according to their business description.

the other European countries according to the data provided in Table 2. There appear to be no significant differences in the geographical distribution of cleantech innovators and cleantech ecosystem firms.

[Insert Table 2 here]

While interest in cleantech has surged in recent years, driven in part by regulatory developments, environmental technology is not a new phenomenon and predates the current green regulatory wave. This is confirmed by our data, as shown in Table 3, which illustrates the distribution of cleantech firms by year of foundation and shows that more than half (60.62%) of the firms were founded before 2000, before the first wave of cleantech investment. Noteworthy is the gradual decline in the number of cleantech firms after 2010. A possible explanation for the initial downward trend is the global economic downturn that occurred during this period, commonly referred to as the "cleantech crash", when the aftermath of the Great Financial Crisis led to reduced investment and financing opportunities for young and innovative firms. These challenges were exacerbated by the sovereign debt crisis, as public budget constraints led to reduced incentives and subsidies for green technologies, creating a less favorable environment for cleantech to flourish. The table also shows a decline from 2021 onwards, which is explained by the methodological approach used to construct the initial sample, which is based on censoring of the most recent firms.

[Insert Table 3 here]

3.4. Control group sample description

To identify a control group of firms from the non-cleantech sample downloaded from Orbis that are more similar to cleantech firms in terms of observable characteristics, we used a two-stage matching procedure. Specifically, we first used coarsened exact matching (CEM, Iacus et al., 2012), followed by propensity score matching (PSM, Rosenbaum and Rubin, 1983).

PSM selects matched firms based on a propensity score, i.e., the probability to "be treated" (in this case, being a cleantech firm) estimated based on a set of matching variables. CEM enforces greater control over the balance of the matched sample because it matches directly on the matching variables, rather than on a combination of them (i.e., the propensity score). The use of CEM followed by PSM combines the advantages of both matching methods.

The matching variables of the CEM are firms' age classes, country (NUTS0 code) and industries (NACE Rev. 2 Codes). The CEM creates strata along all these dimensions and then retains only observations that fall into strata where there are both cleantech and non-cleantech firms. In other words, the CEM ensures that the control group and the group of cleantech firms overlap along all these dimensions simultaneously. The PSM is then used to identify, among the firms selected by the CEM, those with the highest propensity score to be cleantech. We estimated the propensity score with a probit where the dependent variable is 1 for cleantech and 0 for non-cleantech. In terms of matching variables, we included dummies for firm age, industry (NACE Rev. 2 Codes) and country. Based on the results of the probit, we computed a propensity score, and for each cleantech, we selected the 20 non-cleantech firms with the closest propensity score ("nearest neighbours"). We ended up with 125,954 matched non-cleantech firms that entered our control group sample.

3.5. Country level sample

We then built a panel database of cleantech and matched firms, in which all firms are observed every year between 1994 and 2023. We collected information from Orbis on several key financial KPIs, such as turnover and number of employees, to provide a more detailed view of the average size of the cleantech firms in our sample.

In total, given the availability of accounting data, the sample consists of a maximum of 1,115,540 observations on 141,970 firms, of which 23,101 are cleantech firms and the remaining 118,869 are matched non-cleantech firms. More specifically, 198,485 data points relate to 23,101 cleantech firms over the period 1994 to 2022, with an average of 8.59 years of coverage per firm. For non-cleantech firms, our sample consists of 917,055 data points on 118,869 non-cleantech firms over the period 1994 to 2022, with a firm being covered for an average of 7.71 years.

From this sample, we finally derived a panel sample aggregated at the country level, reporting several measures of cleantech firms for each year. The final sample thus consists of 29 European countries observed from 1994 to 2023[3]. All estimated measures were estimated by distinguishing between categories: innovators and all cleantech ecosystem firms (i.e. Manufacturers, Distributors, Operators, Experimenters and Integrators).

More specifically, with regard to the measures used to describe cleantech activity in our European countries over time, we first estimated the percentage of newborn cleantech firms (i.e. firms with a founding year equal to the focal year) out of the total number of newborn firms (i.e. cleantech and non-cleantech). This measure has been used as a proxy of cleantech activity in a specific country and year. As shown in Table 4, 1.43% of newborn firms are cleantech innovators. This percentage increases to 8.090% when considering cleantech ecosystem firms. Looking at the different categories, Operators (2.406%) and Integrators (2.137%) are the most represented categories in the cleantech ecosystem world in terms of percentage of newborn firms.

Second, for each country and year, we also estimated the average growth of newborn cleantech firms over the next three years (i.e. from the focal years to the next two years). We measured growth in terms of both sales (i.e. $logSales_t$-$logSales_{t-1}$) and number of employees (i.e.

[3] Countries included in our final sample are a subsample of the original one since we exclude countries for which we were not able to retrieve any information about policy regulation (as will be described in Section 4). Countries included in the final sample are the following: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Croatia, Hungary, Ireland, Iceland, Italy, Liechtenstein, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia and Slovakia, observed from 1994 to 2023.

$logEmployees_t$-$logEmployees_{t-1}$). These measures are used as proxies for the performance of cleantech firms in a given country and year. As shown in Table 4, newborn cleantech innovators grow at an average annual rate of 12.3% and 11.6%, in terms of sales and employees respectively. This percentage increases to 17.7% when looking at sales growth in cleantech ecosystem firms while, for employment growth, cleantech ecosystem firms show an average percentage (11.1%) very similar to that of cleantech firms. Looking at the different categories, Operators and Integrators (2.137%) are still the categories showing a higher annual performance increase.

[Insert Table 4 here]

4. Policies database construction

As the European Green Deal (EGD) aims to define a broad policy strategy, it sets targets for the development of specific measures in eight key areas: i) raising the EU's climate ambition; ii) providing clean, affordable and secure energy; iii) mobilising industry for a clean and circular economy; iv) building and renovating in an energy and resource efficient way; v) striving for a pollution-free environment; vi) preserving and restoring ecosystems and biodiversity; vii): achieving a fair, healthy and environmentally friendly food system ("Farm to Fork"); viii) accelerating the transition to sustainable and smart mobility.

To build the dataset of policies, we examined the key areas of the EGD for the sample of countries where we were able to identify at least one cleantech firm[4]. Each step of the dataset construction procedure is described in detail below.

[4] The countries included in our final sample are a sub-sample of the original sample, as we excluded countries for which we were unable to retrieve any information on policy regulation (as described in Section 4). Countries included in the final sample are the following: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Croatia, Hungary, Ireland, Iceland, Italy, Liechtenstein, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia and Slovakia, observed from 1994 to 2023.

First, we analysed the different chapters of the EGD[5], together with the communications from the EU Commission on the different topics and compiled a list of keywords representative of each topic covered by the chapters. On the basis of the list of keywords identified, the legislative body of the European Union was searched to identify the most relevant Directives and Regulations contributing to the achievement of the objectives of the EGD, published in the period between 2000 and 2023. Recognizing that Regulations enter into force immediately and are legally binding for each EU Member State once ratified by the EU Parliament, whereas Directives have to be transposed into national law, we proceeded from the higher to the lower administrative level (i.e. starting from the EU legislative body, to the national level, and then to the regional level).

The main source of information was EUR-Lex[6], the EU's information platform, where both EU-level legislation and national transposition can be found. In particular, the "National transposition" section provided information on the measures taken by the Member States to transpose EU legislation into national law. The search was carried out on the basis of the titles and, where available, the texts of the measures. Where the texts were not directly available, they were located on national government websites and online repositories.

Once identified as relevant legislation, each measure was coded and classified. The coding procedure was carried out on the basis of the date of notification, the identifiers of the acts and, in case of national transpositions, the parent directive. The classification of each measure takes several aspects into account. First, the policy level was identified, i.e. EU-level (EU), Country (CO) or Region (RE). Second, the typology of the legislative instrument was defined among several options, namely EU Commission's strategy, Standard definition, Incentives-taxes, Binding Targets for all Member States, Standards and obligations, Administrative

[5] European Commission. (2020, December 11). European Green Deal. Retrieved from https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal
[6] https://eur-lex.europa.eu/homepage.html

arrangement, Infrastructure design, National regulation (Botta and Koźluk, 2014). Third, where available, we identified a performance indicator that forms the basis of the measure, such as the amount of GHG emitted by fuels, reduction in energy intensity, etc. Four, the scope of the measure was identified, i.e. the factor on which the measure is based (e.g. firm size, industrial sector, specific technology etc.). Finally, the existence of amendments or the expiry of the measure was noted, together with the reference dates.

The list of measures was considered exhaustive, if the search yielded only repetitive items. The results of the classification were shared among the research team and internal consistency checks were carried out independently by each member of the team.

Figure 1 summarises the methodology used to identify and classify the relevant policy measures.

[Insert Figure 1 here]

From this database, which collects all information on policy regulation, we derived two main variables, at the country and year level. The first one is a dummy variable, $d\_policies_{t-1}$, which takes the value 1 if the focal country introduced at least one policy regulation until the year before the focal year (*t-1*), and 0 otherwise. The second variable, $n\_cumulative\ policies_{t-1}$, measures the cumulative number of policies introduced by the focal country up to the year before the focal year (*t-1*).

Table 5 reports some descriptive statistics on these policy variables. 72% of the observations in our sample refer to countries that implemented at least one policy regulation in a given year with respect to the year before the focal year. On average, countries implemented about 9 policies from the beginning of the observation period to the focal year.

[Insert Table 5 here]

5. Empirical results

5.1. Descriptive statistics

It is interesting to provide some preliminary descriptive statistics on the impact of the introduction of policy regulation on the cleantech sector.

In Table 6, we compare our measures of cleantech activity (% of newborn firms, sales and employee growth, as described in Section 3.5) between countries that introduced at least one policy (Column I) in year t and countries that did not introduce any regulation in previous years (Column II). We also explore whether interesting differences emerge when going deeper into the categorization of cleantech ecosystem firms.

The descriptive statistics suggest that the birth rate of cleantech innovators is positively affected by the introduction of policies, even though the difference is not significant. It has a positive and significant effect in terms of growth of cleantech innovator firms in the following three years, both in terms of sales and employees. For cleantech ecosystem firms, the introduction of policies per se shows a negative and significant effect on the birth rate of new cleantech ecosystem firms, but a positive effect on the growth of cleantech ecosystem firms, both in terms of sales and employees, is confirmed. These results seem to be consistent across all categories of the cleantech ecosystem world.

In the last columns of Table 6, we examine the effect of the number of cumulative policies introduced by a given country, reporting the average value of our measures of cleantech activity (% of newborn firms, sales and employee growth), according to the four quartiles of the distribution of the number of cumulative policies. The figures in the last four columns of Table 6 suggest that the higher the quartiles of the number of cumulative policies, the lower the percentage of newborn firms. This effect seems to be particularly evident for cleantech ecosystem firms, while the trend is not linear for cleantech innovators.

In terms of growth measures, the results are more heterogeneous: for cleantech innovators, there is no clear trend between the number of cumulative policies and growth, both in terms of sales and employees. For the cleantech ecosystem, the descriptive results suggest an increase in sales growth associated with the introduction of a higher number of policies, while the trend is not clear when looking at employee growth.

[Insert Table 6 here]

[Insert Figure 2 here]

5.2. Regression analyses

To make this analysis more robust, we estimated several econometric models to test our hypotheses. Our first model aims to estimate the impact of policies on the birth rate of cleantech firms. To explore this effect, we used a logit model with the percentage of newborn cleantech firms out of the total number of newborn firms in a given country and year as dependent variable. Among the independent variables, we included the two variables that proxy the effect of policy regulation as described in Section 4. We also included country and year dummies as controls. Table A in the Appendix shows the correlation matrix between all the variables of interest.

The results of this estimates are shown in Table 7. In more detail, we examined this effect by distinguishing between cleantech innovators (I column) and cleantech ecosystem firms (II column).

The results indicate that the introduction of policy regulation in the previous years alone has a positive and significant (at 10% confidence level) effect on the birth rate of cleantech innovator firms, while the number of policies introduced has no significant effect. This seems in accordance with what found in descriptive statistics in Table 6, even though the difference was not significant at standard confidence levels. Conversely, the introduction of policies has no significant impact on cleantech ecosystem firms, while the number of policies introduced has a negative and significant (at 1% confidence level) effect on the cleantech ecosystem firm birth rate. Again, this result confirms what is shown in the descriptive statistics in Table 6.

[Insert Table 7 here]

We also estimated the same model disentangling between cleantech ecosystem firms by the categories defined in Section 3.3. The results of these estimates are reported in Table 8.

Looking at the different categories of cleantech ecosystem firms, there are no significant differences in the number of policies introduced, while the introduction of policies seems to have a positive and significant effect (at 10% confidence level) on Experimenters, similar to what happens for cleantech innovator firms, and a negative and significant effect (at 10% confidence level) for Integrators.

[Insert Table 8 here]

The second model is a regression in which we resorted to the performance measures described in section 3.5 as dependent variables. We also included the average growth of all firms in the focal country and year as a control. The results of these analyses are shown in Table 9. In more detail, we examined this effect by distinguishing between cleantech innovators (I-II columns) and cleantech ecosystem firms (III-IV columns). We used both the average growth rate of sales (columns I and III) and employees (columns II and IV) in the following three years after foundation.

The results suggest that the introduction of policies favours cleantech innovators in terms of employee growth, while the effect on sales is not significant. Looking at cleantech ecosystem firms, again, it is not the introduction of policies per se that influences the growth of cleantech ecosystem firms, but there is a positive and significant impact of the cumulative number of policies introduced on the growth of firms, both in terms of sales and employees, according to the statistics reported in Table 6.

[Insert Table 9 here]

In addition, we estimated the same model disentangling between cleantech ecosystem firms by the categories defined in Section 3.3. The results of these estimates are reported in Table 10 for sample growth and in Table 11 for employee growth. Looking at sales growth, the results discussed for the overall sample of cleantech ecosystem firms are confirmed: all categories seem to be positively affected by the cumulative number of policies introduced in the focal country, except for Distributors. Looking at employee growth, the positive and significant effect of the cumulative number of policies introduced in the focal country is confirmed for Experimenters and Manufacturers, while a positive and slightly significant effect (10% confidence level) is found for Distributors.

[Insert Table 10 here]

Summarizing, our results suggest that the introduction of national cleantech policies may promote the creation and employee's growth of cleantech innovator firms by providing financial support, market incentives, and regulatory guidance, and by making it more attractive for entrepreneurs to enter the cleantech sector. When considering the cleantech ecosystem, our analysis suggest that it is influenced by the number of national cleantech policies introduced in a given country: if, on the one hand, obtained results confirm a positive effect on the growth of existing companies, on the other hand suggests, in accordance with previous evidence, that an excessive number of complex and sometimes overlapping policies can potentially introduce compliance burdens, uncertainty and regulatory risk, market fragmentation, regulatory barriers to entry, or lack of coordination, this resulting in a negative impact on the creation of new cleantech ecosystem companies.

6. Conclusions and Policy Implications

Achieving net-zero emissions by 2050 is the challenge of the current generation on a global scale, and the role that the EU has carved out for itself is one of global leadership, with the protection of human health and well-being, as well as the creation of jobs and economic growth, as its founding principles. For these reasons, the strategy outlined in the EGD must be assessed not only in terms of the environmental protection it ensures, but also in terms of its ability to create a supportive framework for green and just growth, by catalysing public investment and private investment in green innovation. The aim of this paper is to explore the complex relationship between cleantech policy and cleantech entrepreneurship, by building extensive datasets covering cleantech firms on the one hand, and the policies implementing the European Green Deal strategy on the other, over a period from 1994 and 2022.

In particular, the paper focuses on assessing the extent to which policy implementation influences the scope of entrepreneurial opportunities and the creation of new firms. The results obtained suggest that the introduction of sustainability-oriented policies has a positive effect on the creation and employees growth of cleantech innovator firms, acting as a trigger for industrial innovation, regardless of the number of policies implemented. However, when considering the industrial innovation ecosystem, the complexity of the policy framework can act as a barrier to the development of new firms: an increasing number of policies has a regulatory effect in the cleantech ecosystem limiting the number of newborn companies while favouring their growth.

The process of building the datasets has allowed some relevant insights to be gathered, which make it possible to envisage avenues for future research. Firstly, policy development appears uneven across the different EGD chapters: some, such as Renewable Energies policies, are already on track to meet the 2030 target (in this case, to increase the share of renewable energy in the energy mix to 40%); others, such as Energy Efficiency measures, as well as Sustainable

Transport and Circular Economy have been tackled and framed, but are not yet fully completed; finally, some chapters are still in their first stages of development, such as “From Farm to Fork” and Biodiversity. Moreover, cross-cutting policies that could be crucial for accelerating the transitions, such as a carbon border adjustment mechanism and a comprehensive sustainable finance framework, are still under discussion. This uneven distribution of policies necessarily leads to an uneven approach of the market and entrepreneurial efforts towards the different transition areas. To date, no sector-specific evaluations have been carried out, leaving room for future research on this pathway.

Secondly, the study of EGD-related policies has made it possible to identify some patterns in Member States attitudes towards the implementation of EU environmental policy. On the one hand, the multi-level governance framework allows Member States to work on the transposition of Directives through both primary legislation (such as Acts of the Parliament or Presidential Decrees) or secondary regulations or ministerial orders, in order to better comply with their political environment and enforcement capacity. These mechanisms provide flexibility, but they also leave room for partial transposition, delays and a different number of domestic policies implementing, for example, the same Directive. On the other hand, the administrative framework varies from country to country, with more centralised approaches in some cases, e.g. France, where the national government takes the lead in transposing directives, or more decentralised approaches, e.g. federal states like Germany, where regional or local governments play a more significant role.

Finally, Member States are either leaders or laggards, due to a variety of contributing factors influencing the speed, effectiveness, and willingness to adopt and enforce environmental regulations. These factors are linked to historical legacies, both in terms of environmental leadership (e.g. the Netherlands, Denmark, and Sweden) or reluctance (e.g. Eastern European countries) and economic development. These characteristics are thought to support the absorptive capacity of the local ecosystem, and possibly benefit from environmental protection measures, resource availability and exposure to climate-related risks, which could support the public awareness, cultural development and, consequently, political pressure for the implementation of such measures (Jänicke and Wurxle, 2019). This attitude can play an important role in the development of an industrial ecosystem that is solid and proactive enough to seize the opportunities offered by new policies in timely manner, or, conversely, suffer from competition at the EU level, from companies developing in leading countries.

References

Ambec, S., Cohen, M.A., Elgie, S., Lanoie, P. (2013). The Porter Hypothesis at 20: Can environmental regulation enhance innovation and competitiveness? Review of Environmental Economics and Policy, 7(1), 2-22.

Ambec, S., Barla, P. (2006). Can environmental regulations be good for business? An assessment of the Porter hypothesis. Energy Studies Review, 14, 42-62.

Ambrois, M., Butticè, V., Caviggioli, F., Cerulli, G., Croce, A., De Marco, A., Giordano, A., Resce, G., Toschi, L., Ughetto, E., Zinilli, A. (2023). Using machine learning to map the European cleantech sector. EIF working paper 2023/91

Ball, C., Kittler, M. (2019). Removing environmental market failure through support mechanisms: Insights from green start-ups in the British, French and German energy sectors. Small Business Economics, 52, 831-844.

Bi, G.B., Song, W., Zhou, P., Liang, L. (2014). Does environmental regulation affect energy efficiency in China's thermal power generation? Empirical evidence from a slacks-based DEA model. Energy Policy, 66, 537-546.

Blackman, A., Lahiri, B., Pizer, W., Planter, M.R., Pina, C.M. (2010). Voluntary environmental regulation in developing countries: Mexico's clean industry program. Journal of Environmental Economics and Management, 60(3), 182-192.

Black, A., Ramanathan, R. and Nath, P., 2010. Flexible environmental regulations and performance: The moderating role of innovation. Academy of Management Annual Meeting, Montreal, Canada, 6–10 August, 2010.

Botta, E., T. Koźluk (2014), "Measuring Environmental Policy Stringency in OECD Countries: A Composite Index Approach", OECD Economics Department Working Papers, No. 1177, OECD Publishing, Paris.

Brunnermeier, S.B., Cohen, M.A. (2003). Determinants of environmental innovation in US manufacturing industries. Journal of Environmental Economics and Management, 45, 278-293.

Bruton, G. D., Ahlstrom, D., Li, H. L. (2010). Institutional theory and entrepreneurship: Where are we now and where do we need to move in the future? Entrepreneurship Theory and Practice, 34(3), 421-440.

Burer M.J., Wustenhagen R. (2009). Which renewable energy policy is a venture capitalist's best friend? Empirical evidence from a survey of international cleantech investors. Energy Policy, 37, 4997-5006.

Chittenden, F., Kauser, S., Poutziouris, P. (2005). PAYE–NIC compliance costs: Empirical evidence from the UK SME economy. International Small Business Journal, 23(6), 635–656.

Christensen, C.M., McDonald, R., Altman, E.J., Palmer, J.E. (2018). Disruptive innovation: An intellectual history and directions for future research. Journal of Management Studies, 55(7), 1043-1078.

Criscuolo, C., Menon, C. (2015). Environmental policies and risk finance in the green sector: Cross-country evidence. Energy Policy, 83, 38–56.

Cumming, D., Henriques, I., Sadorsky, P. (2016). 'Cleantech' venture capital around the world. International Review of Financial Analysis, 44, 86–97

Eiadat, Y., Kelly, A., Roche, F., Byadat, H. (2008). Green and competitive? An empirical test of the mediating role of environmental innovation strategy. Journal of World Business, 43, 131-145.

European Commission. (2019). The European Green Deal [Online]. Available: https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1&format=PDF

Feinerer, I., Hornik, K., Meyer, D. (2008). Text Mining Infrastructure in R. Journal of Statistical Software, 25(5).
Filbeck, G., Gorman, R.F. (2004). The relationship between the environmental and financial performance of public utilities. Environmental and Resource Economics, 29(2), 137-157.
Ford, J.A., Steen, J., Verreynne, M.L. (2014). How environmental regulations affect innovation in the Australian oil and gas industry: going beyond the Porter Hypothesis. Journal of Cleaner Production, 84, 204-213.
Gray, W.B., Shadbegian, R.J. (2003). Plant vintage, technology, and environmental regulation. Journal of Environmental Economics and Management, 46, 384-402.
Grubb, M. (2004). Technology innovation and climate change policy: An overview of issues and options. Keio Economic Studies, 41(2), 103–132.
Hiatt, S. R., Grandy, J. B., Lee, B. H. (2015). Organizational responses to public and private politics: An analysis of climate change activists and U.S. oil and gas firms. Organization Science, 26(6), 1769-1786.
Horbach, J. (2008). Determinants of environmental innovation: New evidence from German panel data sources. Research Policy, 37, 163-173.
Hu, D., Wang, Y., Huang, J., Huang, H. (2017). How do different innovation forms mediate the relationship between environmental regulation and performance? Journal of Cleaner Production, 161, 466-476.
Iacus S.M., King G., Porro G. (2012). Causal inference without balance checking: Coarsened exact matching. Political Analysis 20, 1 (2012), 1–24.
IPCC. (2014). Climate Change 2014: Impacts, Adaptation and Vulnerability. In W. G. II (Ed.), Fifth Assesment Report (AR5). Geneva, Switzerland: Intergovernmental Panel on Climate Change.
Jaffe, A.B., Palmer, K. (1997). Environmental regulation and innovation: A panel data survey. Review of Economics and Statistics, 79(4), 610-619.
Johnstone, N., Haˇsˇciˇc, I., Kalamova, M. (2010). Environmental policy characteristics and technological innovation. Economia politica, 27(2), 277–302.
Kitching, J. (2006). A burden on business? Reviewing the evidence base on regulation and small business performance. Environment and Planning C: Government and Policy, 24(6), 799–814.
Klapper, L., Laeven, L., Rajan, R. (2006). Entry regulation as a barrier to entrepreneurship. Journal of Financial Economics, 82, 591-629.
Kirkpatrick, C., Parker, D. (2015). Measuring Regulatory Performance. OECD.
Lanoie, P., Laurent-Lucchetti, J., Johnstone, N., Ambec, S. (2011). Environmental policy, innovation and performance: New insights on the Porter hypothesis. Journal of Economics & Management Strategy, 20(3), 803-842.
Lopez-Gamero, M.D., Molina-Azorin, J.F., Claver-Cortes, E. (2010). The potential of environmental regulation to change managerial perception, environmental management, competitiveness, and financial performance. Journal of Cleaner Production, 18(10/11), 963-974.
Majumdar, S.K., & Marcus, A.A. (2001). Rules versus discretion: The productivity consequences of flexible regulation. Academy of Management Journal, 44, 170-179.
Martínez-Zarzoso, I., Bengochea-Morancho, A., Morales-Lage, R. (2019). Does environmental policy stringency foster innovation and productivity in OECD countries? Energy Policy, 134, 110982
Montabon, F., Sroufe, R., Narasimhan, R. (2007). An examination of corporate reporting, environmental management practices, and firm performance. Journal of Operations Management, 25, 998-1014.

Mura, M., Longo, M., Zanni, S., Toschi, L. (2023). Exploring socio-economic externalities of development scenarios. An analysis of EU regions from 2008 to 2016. Journal of Environmental Management, 332, 117327.

Pacheco, D.F., Dean, T.J., Payne, D.S. (2010). Escaping the green prison: entrepreneurship and the creation of opportunities for sustainable development. Journal of Business Venturing, 25, 464-480.

Pernick, R., & Wilder, C. (2007). The clean tech revolution: The next big growth and investment opportunity. Collins

Porter, M.E., van der Linde, C. (1995). Green and competitive: ending the stalemate. Harvard Business Review, 73(5), 120-134.

Porter, M.E. (1991). America's green strategy. Scientific American, 264(4), 168.

Porter, M.E., Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. Journal of Economic Perspectives, 9(4), 97-118.

Ramanathan, R., He,Q., Black, A., Ghobadian, A., Gallear, D. (2017). Environmental regulations, innovation and firm performance: A revisit of the Porter hypothesis. Journal of Cleaner Production, 155(2), 79-92.

Ramanathan, R., Black, A., Nath, P., Muyldermans, L. (2010). Impact of environmental regulations on innovation and performance in the UK industrial sector. Management Decision, 48(10), 1493-1513.

Ramus C. A., Montiel I. (2005). When are corporate environmental policies a form of greenwashing? Business & Society, 44, 377-414.

Rennings K., Rammer, C. (2011). The impact of regulation-driven environmental innovation on innovation success and firm performance. Industry and Innovation, 18(3), 255-283.

Rosenbaum P.R., Rubin D.B. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rubashkina, Y., Galeotti, M., Verdolini, E. (2015). Environmental regulation and competitiveness: Empirical evidence on the Porter Hypothesis from European manufacturing sectors. Energy Policy, 83, 288-300.

Sanchez, C.M., McKinley, W. (1998). Environmental regulatory influence and product innovation: The contingency effects of organizational characteristics. Journal of Engineering and Technology Management, 15(4), 257-278.

Sunny, S.A., Shu, C. (2017). Investments, incentives, and innovation: geographical clustering dynamics as drivers of sustainable entrepreneurship. Small Business Economics, 1-23.

Shao, S., Hu, Z., Cao, J., Yang, L., Guan, D. (2020). Environmental Regulation and Enterprise Innovation: A Review. Business Strategy and the Environment, 29(8), 3479-3493.

Tello, S.F., Yoon, E. (2008). Examining drivers of sustainable innovation. Journal of International Business Strategy, 8(3), 164-169.

Triebswetter, U., Hitchens, D. (2005). The impact of environmental regulation on competitiveness in the German manufacturing industry: A comparison with other countries in the European Union. Journal of Cleaner Production, 13, 733-745.

Triebswetter, U., Wackerbauer, J. (2008). Integrated environmental product innovation in the region of Munich and its impact on company competitiveness. Journal of Cleaner Production, 16(14), 1484-1493.

Yang, Y., Liu, X. (1999). A re-examination of text categorization methods. 42–49

York, J. G., Lenox, M. J. (2014). Exploring the sociocultural determinants of de novo versus de alio entry in emerging industries. Strategic Management Journal, 35, 1930–1951.

Walker K., Wan F. (2012). The harm of symbolic actions and green-washing: Corporate actions and communications on environmental performance and their financial implications. Journal of Business Ethics, 109, 227-242.

Tables

Table 1: Classification of cleantech firms into different ecosystem segments

	n.comp	*%*
Cleantech innovators	3,201	13.30%
Cleantech ecosystem	20,867	86.70%
Experimenters	103	0,43%
Manufacturers	5,380	22.35%
Distributors	3,337	13.86%
Integrators	6,558	27.25%
Operators	5,489	22.81%
Total	*24,068*	*100%*

Table 2: Distribution of Cleantech firms by country

	Cleantech innovators		*Cleantech ecosystem*	
Country ISO code	*n.firms*	*%*	*n.firms*	*%*
AL	1	0.031%	0	0.000%
AT	94	2.937%	539	2.467%
BA	0	0.000%	2	0.009%
BE	108	3.374%	635	2.906%
BG	29	0.906%	304	1.391%
CH	3	0.094%	39	0.178%
CY	1	0.031%	2	0.009%
CZ	104	3.249%	665	3.043%
DE	558	17.432%	4058	18.570%
DK	60	1.874%	313	1.432%
EE	14	0.437%	73	0.334%
ES	341	10.653%	1840	8.420%
FI	74	2.312%	463	2.119%
FR	382	11.934%	3146	14.397%
GB	79	2.468%	324	1.483%
GR	46	1.437%	194	0.888%
HR	24	0.750%	170	0.778%
HU	35	1.093%	395	1.808%
IE	2	0.062%	27	0.124%
IS	2	0.062%	16	0.073%
IT	568	17.744%	3,730	17.069%
LT	18	0.562%	135	0.618%
LU	9	0.281%	52	0.238%
LV	5	0.156%	107	0.490%
ME	0	0.000%	12	0.055%
MK	2	0.062%	43	0.197%
MT	3	0.094%	11	0.050%
NL	71	2.218%	347	1.588%
NO	90	2.812%	626	2.865%
PL	158	4.936%	1,312	6.004%
PT	47	1.468%	415	1.899%
RO	47	1.468%	511	2.338%
RS	18	0.562%	222	1.016%
SE	148	4.624%	733	3.354%
SI	24	0.750%	126	0.577%
SK	31	0.968%	242	1.107%
TR	5	0.156%	14	0.064%
Total	*3,201*	*100.000%*	*21,843*	*100.000%*

Table 3: Distribution of Cleantech firms by year of incorporation

Year of incorporation	*Cleantech innovators*		*Cleantech ecosystem*	
	n.firms	*%*	*n.firms*	*%*
Before 1980	530	16.56%	3,777	17.29%
1981-1985	140	4.37%	1,201	5.50%
1986-1990	256	8.00%	2,052	9.40%
1991-1995	351	10.97%	3,451	15.80%
1996-2000	414	12.93%	3,007	13.77%
2001-2005	433	13.53%	2,858	13.09%
2006-2010	563	17.59%	2,753	12.61%
2011-2015	251	7.84%	1,423	6.52%
2016-2020	243	7.59%	1,221	5.59%
2021 onwards	20	0.62%	96	0.44%
Total	*3,201*	*100%*	*21,839*	*100%*

Table 4: Descriptive statistics on cleantech variables

% of newborn cleantech firms		*Mean*	*St.Dev.*	*Min*	*Max*
Cleantech innovators		1.430%	0.041	0	1.000
Cleantech ecosystem		8.090%	0.082	0	0.500
	Distributors	1.290%	0.017	0	0.160
	Experimenters	0.035%	0.001	0	0.014
	Integrators	2.137%	0.028	0	0.200
	Manufacturers	1.726%	0.027	0	0.500
	Operators	2.406%	0.036	0	0.500
Sales growth ($logSales_t$-$logSales_{t-1}$) in the following 3 years		*Mean*	*St.Dev.*	*Min*	*Max*
Average sales growth		0.124	0.247	-1.126	2.041
Cleantech innovators		0.123	0.518	-4.562	3.881
Cleantech ecosystem		0.177	0.335	-0.872	3.169
	Distributors	0.106	0.337	-2.078	2.243
	Experimenters	0.006	0.058	-0.325	0.354
	Integrators	0.136	0.325	-1.198	2.494
	Manufacturers	0.117	0.390	-1.799	3.169
	Operators	0.180	0.467	-1.994	4.062
Employees growth ($logEmp_t$-$logEmp_{t-1}$) in the following 3 years		*Mean*	*St.Dev.*	*Min*	*Max*
Average employees growth		0.079	0.195	-1.079	1.385
Cleantech innovators		0.116	0.412	-0.816	2.471
Cleantech ecosystem		0.111	0.264	-0.305	4.043
	Distributors	0.060	0.318	-0.776	6.906
	Experimenters	0.003	0.033	-0.059	0.384
	Integrators	0.107	0.321	-0.693	4.043
	Manufacturers	0.087	0.281	-0.311	2.428
	Operators	0.074	0.239	-1.631	1.922

Table 5: Descriptive statistics on policy variables

	Mean	St.Dev.	Min	Max
n. cumulative policies	9.096	13.213	0	79
d_policies	0.721	0448	0	1

Table 6: Preliminary evidence on the impact of policies on cleantech activity

	Countries implementing at least one policy d_policies$_t$=1	*Countries non implementing any policy* d_policies$_t$=0	*Diff.*	*Sig.*	*Quartiles of n. policies* *1°* n=0	*2°* n<=3	*3°* 3<n <=12	*4°* n>12
% of newborn cleantech firms								
Cleantech innovators	14.400%	14.000%	0.400%		1.40%	2.38%	1.39%	0.56%
Cleantech ecosystem	7.000%	10.800%	-3.800%	***	10.82%	9.44%	7.38%	4.30%
Distributors	1.100%	1.800%	-0.700%	***	1.85%	1.45%	1.15%	0.62%
Experimenters	0.027%	0.054%	-0.027%	**	0.05%	0.06%	0.02%	0.01%
Integrators	1.820%	2.956%	-1.136%	***	2.96%	2.57%	1.97%	0.92%
Manufacturers	1.467%	2.390%	-0.923%	***	2.39%	2.18%	1.73%	0.51%
Operators	1.957%	3.568%	-1.611%	***	3.56%	2.88%	2.13%	0.88%
Sales growth (logSales$_t$-logSales$_{t-1}$) in the following 3 years								
Cleantech innovators	0.146	0.064	0.082	**	0.064	0.252	0.172	0.018
Cleantech ecosystem	0.230	0.039	0.190	***	0.039	0.117	0.247	0.326
Distributors	0.139	0.020	0.119	***	0.020	0.074	0.151	0.193
Experimenters	0.008	0.000	0.009	*	0.000	-0.005	0.001	0.028
Integrators	0.189	-0.001	0.189	***	-0.001	0.088	0.221	0.261
Manufacturers	0.156	0.017	0.139	***	0.016	0.057	0.167	0.245
Operators	0.242	0.019	0.223	***	0.019	0.126	0.249	0.352
Employees growth (logEmp$_t$-logEmp$_{t-1}$) in the following 3 years								
Cleantech innovators	0.143	0.049	0.094	***	0.048	0.303	0.096	0.030
Cleantech ecosystem	0.148	0.015	0.133	***	0.015	0.136	0.126	0.182
Distributors	0.084	-0.001	0.084	***	-0.001	0.057	0.062	0.131
Experimenters	0.004	-0.001	0.005	*	0.000	-0.001	0.003	0.011

Integrators	0.145	0.008	0.137	***	0.008	0.165	0.091	0.177
Manufacturers	0.115	0.013	0.103	***	0.012	0.124	0.102	0.120
Operators	0.099	0.007	0.092	***	0.007	0.071	0.140	0.090

Table 7. Estimation results: percentage of newborn cleantech firms

	% of newborn cleantech firms			
	Cleantech Innovators		Cleantech Ecosystem	
n_cumulative policies$_{t-1}$	-0.016		-0.220	***
	(0.022)	*[0.482]*	(0.059)	*[0.000]*
d_policies$_{t-1}$	0.764	*	-0.443	
	(0.451)	*[0.090]*	(0.690)	*[0.521]*
Const.	-2.848	**	1.060	
	(1.107)	*[0.010]*	(2.392)	*[0.658]*
Year Fixed Effects	Yes		Yes	
Country Fixed Effects	Yes		Yes	
N. Observations	675		728	

Note: The table reports the logit estimates. The dependent variable is the percentage of newborn cleantech firms out of the total number of newborn firms in a given country and year. For the sake of brevity, we do not report the estimated coefficients for the year and country dummies. Robust standard errors in round brackets; p-values in italics and square brackets. Coefficients and standard errors have been rounded to three decimal places. Significance at 1% level***, 5% level ** and 10% level*.

Table 8. Estimation results: percentage of newborn firms by category of cleantech ecosystem

	% of newborn cleantech firms									
	Distributors		Experimenters		Integrators		Manufacturers		Operators	
n_cumulative policies$_{t-1}$	-0.014		0.077		-0.031		0.005		-0.003	
	(0.029)	*[0.628]*	(0.051)	*[0.130]*	(0.033)	*[0.351]*	(0.028)	*[0.851]*	(0.033)	*[0.931]*
d_policies$_{t-1}$	-0.557		1.173	*	-1.034	*	-0.046		0.097	
	(0.492)	*[0.257]*	(0.671)	*[0.080]*	(0.544)	*[0.057]*	(0.489)	*[0.926]*	(0.589)	*[0.870]*
Const.	-3.709	**	-5.714	***	1.702		-4.634	***	-3.349	***
	(1.734)	*[0.032]*	(1.835)	*[0.002]*	(1.182)	*[0.150]*	(1.643)	*[0.005]*	(1.273)	*[0.009]*
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes	
Country Fixed Effects	Yes		Yes		Yes		Yes		Yes	
N. Observations	676		336		368		729		504	

Note: The table reports the logit estimates. The dependent variable is the percentage of newborn cleantech firms out of the total number of newborn firms in a given country and year by category of cleantech ecosystem. For the sake of brevity, we do not report the estimated coefficients for the year and country dummies. Robust standard errors in round brackets; p-values in italics and square brackets. Coefficients and standard errors have been rounded to three decimal places. Significance at 1% level***, 5% level ** and 10% level*.

Table 9. Estimation results: newborn firm performance

	Cleantech innovators				Cleantech Ecosystem			
	Sales growth		Employees growth		Sales growth		Employees growth	
n_cumulative policies$_{t-1}$	-0.004		-0.003		0.009	***	0.003	**
	(0.003)	*[0.260]*	(0.002)	*[0.112]*	(0.002)	*[0.000]*	(0.002)	*[0.038]*
d_policies$_{t-1}$	0.068		0.079	**	0.01		0.009	
	(0.069)	*[0.327]*	(0.038)	*[0.040]*	(0.036)	*[0.777]*	(0.033)	*[0.779]*
Average performance	0.328	***	0.585	***	0.383	***	0.375	***
	(0.082)	*[0.000]*	(0.060)	*[0.000]*	(0.043)	*[0.000]*	(0.053)	*[0.000]*
Const.	0.01		-0.076		-0.375	***	-0.077	
	(0.162)	*[0.949]*	(0.133)	*[0.569]*	(0.085)	*[0.000]*	(0.116)	*[0.507]*
Year Fixed Effects	Yes		Yes		Yes		Yes	
Country Fixed Effects	Yes		Yes		Yes		Yes	
N. Observations	812		812		812		812	

Note: The table reports the OLS estimates. The dependent variable is average growth (sales and employees) in the following three years of newborn cleantech firms, distinguishing among innovators and ecosystem. For the sake of brevity, we do not report the estimated coefficients for the year and country dummies. Robust standard errors in round brackets; p-values in italics and square brackets. Coefficients and standard errors have been rounded to three decimal places. Significance at 1% level***, 5% level ** and 10% level*.

Table 10. Estimation results: newborn firm performance (sales growth) by category of cleantech ecosystem firms

	Sales growth									
	Distributors		Experimenters		Integrators		Manufacturers		Operators	
n_cumulative policies$_{t-1}$	0.001		0.001	***	0.008	***	0.004	**	0.01	***
	(0.002)	*[0.514]*	(0.000)	*[0.001]*	(0.002)	*[0.000]*	(0.002)	*[0.048]*	(0.003)	*[0.000]*
d_policies$_{t-1}$	-0.029		-0.005		0.002		-0.046		0.013	
	(0.042)	*[0.492]*	(0.007)	*[0.495]*	(0.038)	*[0.948]*	(0.048)	*[0.341]*	(0.058)	*[0.820]*
Average sales growth$_t$	0.092	*	0.003		0.145	***	0.167	***	0.327	***
	(0.049)	*[0.064]*	(0.008)	*[0.708]*	(0.045)	*[0.001]*	(0.057)	*[0.003]*	(0.068)	*[0.000]*
Const.	-0.077		-0.029	*	-0.33	***	-0.190	*	-0.417	***
	(0.098)	*[0.432]*	(0.017)	*[0.079]*	(0.09)	*[0.000]*	(0.113)	*[0.092]*	(0.135)	*[0.002]*
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes	
Country Fixed Effects	Yes		Yes		Yes		Yes		Yes	
N. Observations	812		812		812		812		812	

Note: The table reports the OLS estimates. The dependent variable is average sales growth in the following three years of newborn cleantech firms, distinguishing among innovators and ecosystem. For the sake of brevity, we do not report the estimated coefficients for the year and country dummies. Robust standard errors in round brackets; p-values in italics and square brackets. Coefficients and standard errors have been rounded to three decimal places. Significance at 1% level***, 5% level ** and 10% level*.

Table 11. Estimation results: newborn firm performance (employees growth) by category of cleantech ecosystem firms

	Employees growth									
	Distributors		Experimenters		Integrators		Manufacturers		Operators	
n_cumulative policies$_{t-1}$	-0.004	*	0.001	***	0.001		0.007	***	-0.001	
	(0.002)	*[0.081]*	(0.000)	*[0.004]*	(0.002)	*[0.785]*	(0.002)	*[0]*	(0.002)	*[0.584]*
d_policies$_{t-1}$	-0.069		-0.005		-0.048		0.021		0.042	
	(0.045)	*[0.12]*	(0.004)	*[0.305]*	(0.039)	*[0.217]*	(0.034)	*[0.531]*	(0.033)	*[0.2]*
Average employees growth$_t$	0.067		-0.002		0.27	***	0.026		0.053	
	(0.071)	*[0.341]*	(0.007)	*[0.754]*	(0.062)	*[0]*	(0.054)	*[0.628]*	(0.052)	*[0.299]*
Const.	0.645	***	0.026	*	0.092		-0.335	***	-0.018	
	(0.156)	*[0]*	(0.015)	*[0.096]*	(0.136)	*[0.499]*	(0.119)	*[0.005]*	(0.114)	*[0.871]*
Year Fixed Effects	Yes		Yes		Yes		Yes		Yes	
Country Fixed Effects	Yes		Yes		Yes		Yes		Yes	
N. Observations	812		812		812		812		812	

Note: The table reports the OLS estimates. The dependent variable is average employees growth in the following three years of newborn cleantech firms, distinguishing among innovators and ecosystem. For the sake of brevity, we do not report the estimated coefficients for the year and country dummies. Robust standard errors in round brackets; p-values in italics and square brackets. Coefficients and standard errors have been rounded to three decimal places. Significance at 1% level***, 5% level ** and 10% level*.

Figures

Figure 3. Methodology followed for the identification and classification of the relevant policy measures

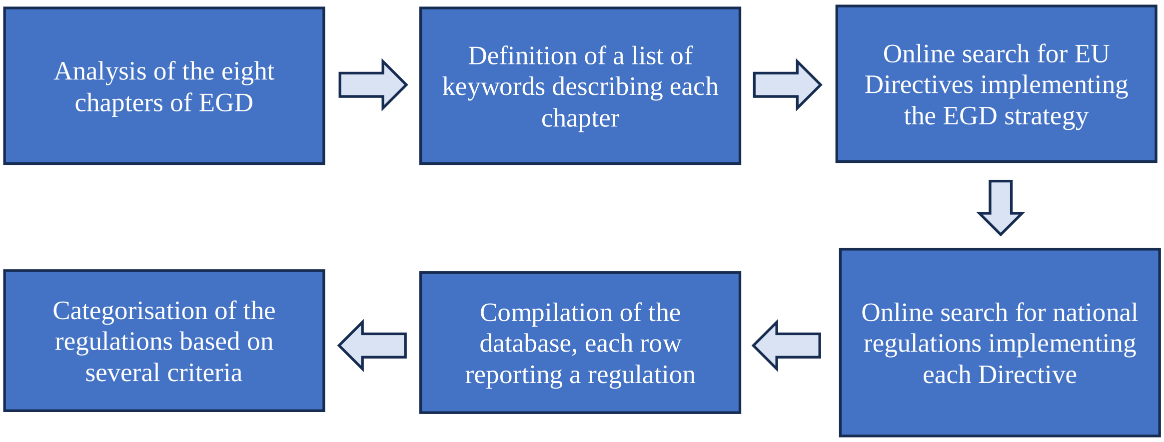
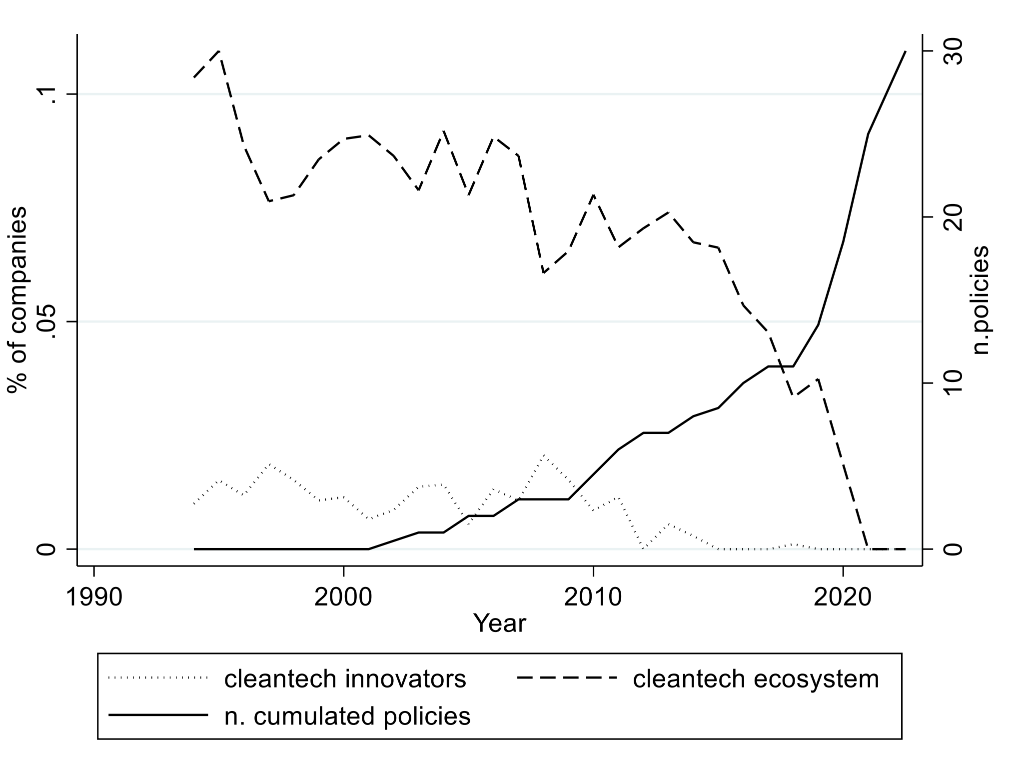


Figure 2. Trend over time: number of cumulative policies and new cleantech firms



Appendix

Table A: Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	% cleantech innovator	1.000																								
2	% cleantech ecosystem	0.157	1.000																							
3	% Distributors	0.136	0.630	1.000																						
4	% Experimenters	0.018	0.170	0.104	1.000																					
5	% Integrators	0.168	0.764	0.507	0.175	1.000																				
6	% Manufacturers	0.102	0.662	0.340	0.136	0.441	1.000																			
7	% Operators	0.106	0.749	0.391	0.084	0.460	0.317	1.000																		
8	Average sales growth	0.064	-0.014	-0.050	0.033	0.031	0.034	-0.029	1.000																	
9	Sales growth cleantech innovator	0.070	0.158	0.086	0.034	0.170	0.079	0.113	-0.008	1.000																
10	Sales growth cleantech ecosystem	-0.005	0.028	-0.009	-0.061	0.061	-0.047	-0.005	-0.095	0.060	1.000															
11	Sales growth Distributors	0.025	0.104	0.049	-0.057	0.068	0.024	0.081	-0.082	0.067	0.384	1.000														
12	Sales growth Experimenters	-0.017	-0.028	-0.028	0.003	0.010	-0.011	-0.040	-0.026	-0.292	0.164	0.079	1.000													
13	Sales growth Integrators	0.001	0.028	-0.009	-0.005	0.070	-0.027	0.004	-0.100	0.131	0.577	0.247	0.123	1.000												
14	Sales growth Manufacturers	-0.033	-0.013	-0.085	-0.006	0.000	-0.028	-0.022	-0.047	0.150	0.544	0.261	0.079	0.280	1.000											
15	Sales growth Operators	-0.008	-0.017	0.006	-0.060	0.033	-0.088	-0.022	-0.084	-0.044	0.678	0.138	0.165	0.426	0.151	1.000										
16	Average Employees growth	0.004	0.111	0.032	0.066	0.129	0.080	0.054	-0.127	0.107	0.052	0.060	0.005	0.062	0.033	0.075	1.000									
17	Employees growth cleantech innovator	-0.027	0.079	0.011	0.045	0.092	0.069	0.046	-0.096	0.168	-0.007	0.101	-0.019	0.013	0.032	-0.017	0.552	1.000								
18	Employees growth cleantech ecosystem	-0.004	0.031	-0.044	-0.043	0.025	-0.009	-0.031	-0.077	0.085	0.388	0.227	0.071	0.242	0.309	0.214	0.350	0.283	1.000							
19	Employees growth Distributors	0.005	-0.015	-0.017	-0.011	0.017	-0.005	-0.033	-0.050	0.096	0.195	0.237	0.063	0.206	0.205	0.152	0.066	0.114	0.288	1.000						
20	Employees growth Experimenters	-0.022	-0.038	-0.040	0.044	-0.015	-0.017	-0.038	-0.020	-0.082	0.075	0.023	0.512	0.078	0.016	0.090	-0.009	-0.058	0.058	0.017	1.000					
21	Employees growth Integrators	0.020	0.034	-0.033	-0.030	0.007	-0.003	-0.017	-0.079	0.089	0.200	0.200	0.046	0.258	0.169	0.156	0.326	0.475	0.701	0.151	0.061	1.000				
22	Employees growth Manufacturers	-0.007	0.002	-0.052	-0.021	0.035	0.000	0.002	-0.056	0.163	0.375	0.269	0.034	0.164	0.603	0.064	0.145	0.039	0.415	0.130	0.038	0.113	1.000			
23	Employees growth Operators	0.002	0.043	0.040	-0.005	0.076	-0.013	0.019	-0.061	-0.011	0.269	0.097	0.135	0.282	-0.010	0.461	0.086	-0.001	0.319	0.058	0.093	0.077	-0.083	1.000		
24	n. cumulative policies	-0.113	-0.267	-0.247	-0.108	-0.247	-0.245	-0.240	-0.133	-0.054	0.264	0.110	0.202	0.260	0.194	0.199	-0.022	-0.108	0.135	0.081	0.183	0.085	0.109	0.035	1.000	
25	d_policies	0.004	-0.207	-0.206	-0.079	-0.177	-0.151	-0.199	-0.015	0.071	0.255	0.158	0.066	0.262	0.160	0.214	0.084	0.102	0.227	0.119	0.066	0.191	0.164	0.173	0.428	1.000

Acknowledgments

We acknowledge the financial support of the European Investment Bank under the EIBURS grant “The cleantech industry in the European Green Deal: policy challenges and the finance landscape for SMEs (CLEU)”