The Impact of Environmental Measures on Trade and Innovation: Evidence from the WTO Environmental Database (EDB)*

Francesco S. Bellelli,† Ankai Xu‡

13th August 2024

Abstract

This study explores the impact of environmental policies on trade and innovation, utilizing an expanded dataset sourced from the WTO Environmental Database (EDB). The study makes two contributions to the literature: First, employing advanced text analysis algorithms, we extracted valuable information from the EDB, enhancing its usefulness for future research and policy analysis. Second, we leverage this augmented dataset to evaluate the impact of environmental measures on trade and green innovation. Our empirical findings suggest that environmental policies not only elevate demand for environmental goods and stimulate imports, but also bolster competitiveness and amplify exports of environmental goods.

JEL classification: F14, F18, O38, Q55, Q58

Keywords: trade and environment; environmental policies; innovation; climate change mitigation

*The opinions expressed in these papers are those of the authors. They do not represent the positions or opinions of the WTO or its members and are without prejudice to members' rights and obligations under the WTO. Any errors are attributable to the authors.

†Quantitative analyst, UBS, Email: F.S.Bellelli@gmail.com

‡Research economist, WTO, Email: ankai.xu@wto.org Address: Rue de Lausanne 154, 1202 Genève, Switzerland. ORCID: 0000-0001-7132-2742. We thank Aik Hoe Lim, Robert Koopman, Sajal Marthur, José-Antonio Monteiro and Daniel Ramos for their insights on the Environmental Database and for providing valuable comments on our work. Additionally, we extend our appreciation to the participants of the 2021 Geneva Trade and Development Workshop, the 2022 Annual Conference of the European Association of Environmental and Resource Economists (EAERE), and the ninth IMF-WBG-WTO Trade Conference for their thoughtful feedback. Any remaining errors remain our responsibility.

1 Introduction

Some of the world's most pressing environmental challenges — such as climate change, deforestation, plastic pollution, and biodiversity loss — demand urgent and innovative responses. In recent years, governments worldwide have implemented a variety of policies aimed at addressing market failures associated with these environmental challenges. Understanding the impact of these policies, particularly in a globalized world, is essential to facilitate a successful transition to a more sustainable future.

This paper leverages the WTO environmental database (EDB) to evaluate the impact of environmental measures on trade and green innovation. The EDB is a repository of over 13,000 environment-related measures reported to the WTO, describing their sectoral focus, environmental objectives, and the types of policy instruments employed. Building upon this existing dataset, our study introduces several extensions tailored to facilitate economic inquiry.[1] Notably, these extensions include: 1) extracting information on the implementation date of the measures, and 2) linking each measure to HS chapters. Leveraging this expanded dataset, we explore the relationship between environmental policies, trade in environmental goods and green innovation. Additionally, we calculate a similarity score that can be used to identify repeating measures in the database and propose a policy strength score that captures both the breadth of a policy measure and the depth of its impact.

To contextualize our study, a number of theoretical and empirical studies in the literature are worth noting. The theoretical literature on environmental policy's impact primarily centers on directed technical change and endogenous growth models (e.g. Acemoglu et al., 2012; Burghaus & Funk, 2013; Acemoglu et al., 2014; Greaker et al., 2018; Hart, 2019; Stöckl, 2020). The seminal work of Acemoglu et al. (2012) emphasizes that, in closed economies, government interventions like R&D subsidies and pollution taxes influence innovation by altering market dynamics. They suggest that a combination of pollution taxation and green R&D subsidies, particularly with early intervention, fosters green innovation under specific circumstances. Similar results are obtained when allowing substitutability-enhancing innovation (Stöckl, 2020), different structure of the innovation market (Greaker et al., 2018) and with different modelling assumptions for the effects of environmental externalities (Burghaus & Funk, 2013; Acemoglu et al., 2016; Hart, 2019).

Several studies also explore the impact of environmental policies in an open economy context. The literature emphasizes that, without international coordination of environmental policies, local environmental regulations could prompt the substitution of locally produced dirty goods with imports (Copeland & Taylor, 2004; Babiker, 2005; Levinson & Taylor, 2008). This pollution haven effect could render domestic production of green goods uncompetitive compared to the dirty imports, thus hampering the green transition (Acemoglu et al., 2014; Hémous, 2016). Broadly similar results are also obtained when assuming that South's research pushes the technological frontier (Di Maria & Smulders, 2004; Hémous, 2016), with different modelling of the innovation market (Witajewski-Baltvilks & Fischer, 2018), by introducing climate feedback effects on capital stocks and different types of interstate policy interactions (Bretschger & Suphaphiphat, 2014).

Empirical research has delved into the impact of environmental policies on innovation and trade patterns. Several studies demonstrate that environmental policies can stimulate innovation in clean technologies across various industries, including the auto industry (Aghion et al., 2012),

[1] The latest version of this extended dataset is available at https://fbellelli.com/EDB-data.html.

wind turbines (Dechezleprêtre & Glachant, 2014) and photovoltaic technology (Peters et al., 2012). Firm-level data also indicate that the EU Emission Trading System stimulates low-carbon innovation (Calel & Dechezleprêtre, 2016). These findings align with the broader literature on the Porter hypothesis, which suggests that well-designed environmental regulations foster innovation and potentially enhance firm competitiveness (Ambec et al., 2013; Dechezleprêtre & Sato, 2017). Empirical work on environmental policies' impact on trade is less conclusive, with some evidence supporting a pollution haven effect, i.e., implementation of more stringent environmental policies leads to a comparative disadvantage in polluting sectors (e.g. Levinson & Taylor, 2008; Kellenberg, 2009; Millimet & Roy, 2016; Koźluk & Timiliotis, 2016; Duan et al., 2021). Few studies, however, have examined environmental policies, innovation, and trade within a unified framework. One exception is Pugliese et al. (2019) that examines scientific, technological and production activities (evidenced by countries' exports) together, in order to identify which capabilities and prerequisites are needed to be competitive in a given activity. The authors find that technology is the best predictor for industrial and scientific production.

Our research makes three main contributions to the literature. First and foremost, we expand upon a distinctive dataset of trade-related environmental measures by employing advanced text analysis algorithms, thereby enhancing its suitability for economic inquiry. This enriched dataset enables us and fellow researchers to conduct a detailed examination of sector-specific implications of environmental policies. Secondly, leveraging the extended data, we provide evidence on the impact of environmental policies on patterns of trade in environmental goods and green innovation. In comparison to studies primarily focused on country-level policy measures and outcomes (e.g. Eugster, 2021; Bettarelli et al., 2023), the matching of environmental measures with broad economic sectors in the extended EDB enables us to explore impacts at a more disaggregated sectoral level, offering a more precise account of their effects. Thirdly, our analysis encompasses a broad spectrum of trade-related policy instruments and provides wider coverage of both developed and developing economies, thus providing a holistic perspective that complements the existing literature, which tends to focus on specific countries or individual policy instruments.

The rest of the paper is structured as follows. Section 2 describes the WTO Environmental Database (EDB) and the steps we take to extend the database for economic research. We present the data and methodology of the empirical analysis in Section 3. Finally, Section 4 concludes.

2 The WTO Environmental Database (EDB)

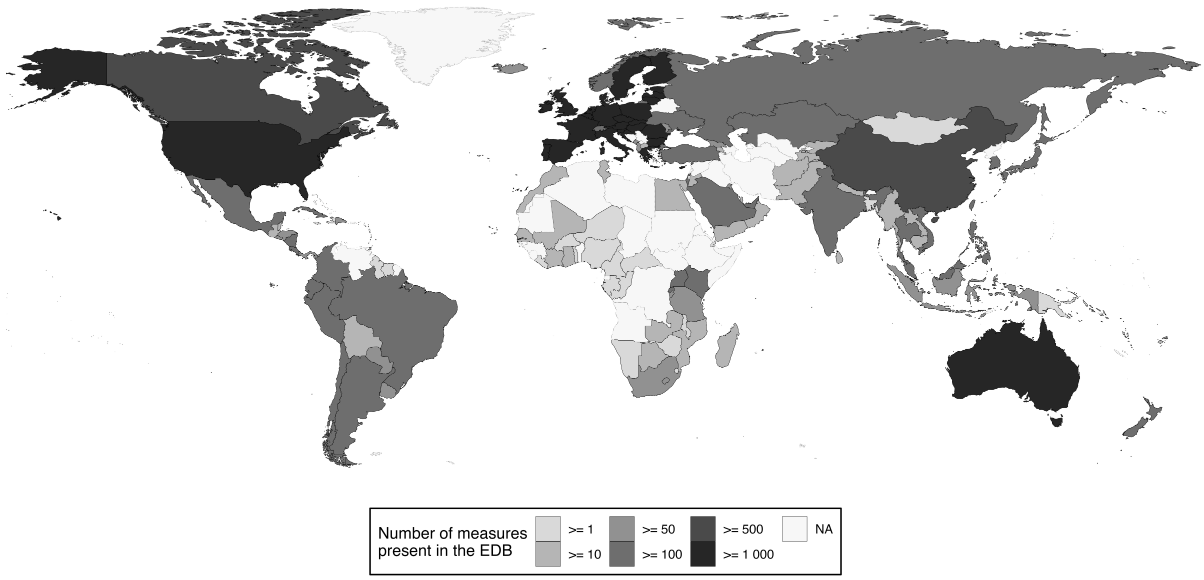
The Environmental Database (EDB)[2] is a collection of information pertaining to trade-related environmental policies of WTO members. The paper primarily draws on a dataset compiled from policies notified by WTO members under multiple WTO agreements. The measures captured in the dataset originate from notifications submitted under the Agreements on Technical Barriers to Trade, Subsidies and Countervailing Measures, Agriculture, Sanitary and Phytosanitary Measures, and Import Licensing Procedures — which together account for nearly 90% of all measures cataloged within the EDB.

[2] Available at: https://edb.wto.org/.

2.1 A unique environmental policy database

The EDB has several advantages over alternative environmental policy databases. First, it exclusively comprises policies measures that bear direct or indirect implications for trade. As a result, while it may not encompass the entirety of a country's environmental policy landscape, it serves as an invaluable resource for research focused specifically on trade-related impacts of environmental policies.

Figure 1: *Number of notified measures by country*

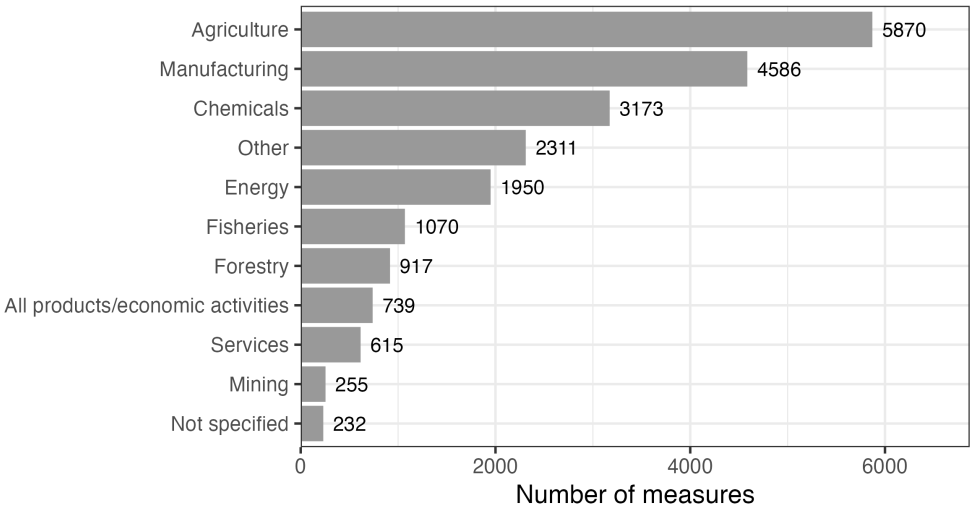


Notes: The map displays the number of notified measures by country. A darker filling indicates that a larger number of measures were notified. Countries and territories in white are those for which the EDB contains no notified measures.

Second, the EDB has an exceptionally broad country coverage. Figure 1 illustrates the number of environment-related measures reported by WTO Members from 2009 to 2022. Unlike other environmental policy databases that often focus primarily on OECD countries or developed nations (e.g. OECD, 2020, 2021; LSE/Columbia Law School, 2021; European Commission, 2021; IRENA, 2021), the EDB has a better representation of developing countries. Furthermore, the EDB offers the added advantage of longitudinal data on environmental policies, enabling the tracking of policy measures over time.

Third, the EDB contains extensive information on the attributes of policy measures. In addition to providing a description for each measure, the EDB readily provides accessible details about the policy instruments of each measure, the specific economic sector to which these policies are applied, and in some instances it includes Harmonized System (HS) and/or International Classification for Standards (ICS) codes for the goods affected by the measure. This makes it ideal for deriving insights on policy design and studying environmental policies at a sectoral level. Notably, agriculture, manufacturing, and the chemicals industry are the most frequently targeted sectors, as shown in Figure 2. Furthermore, the database covers various policy instrument types, including technical regulations, taxes, grants, loans, tariffs, intellectual property measures, export quotas, and non-monetary support or tax concessions.

Figure 2: *Number of measures by sector*



Notes: Some measures relate to two or more sectors.

2.2 Extending the EDB for economic research

Although the EDB contains a wealth of information, certain data are presented in textual formats, which limits their immediate usability for quantitative analyses. A primary objective of this paper is to enhance the accessibility of the EDB for researchers by extracting valuable information that can be applied to economic research. We provide below a succinct overview of our contributions in this regard. For the interested reader, technical details on the extraction process are provided in the appendices.

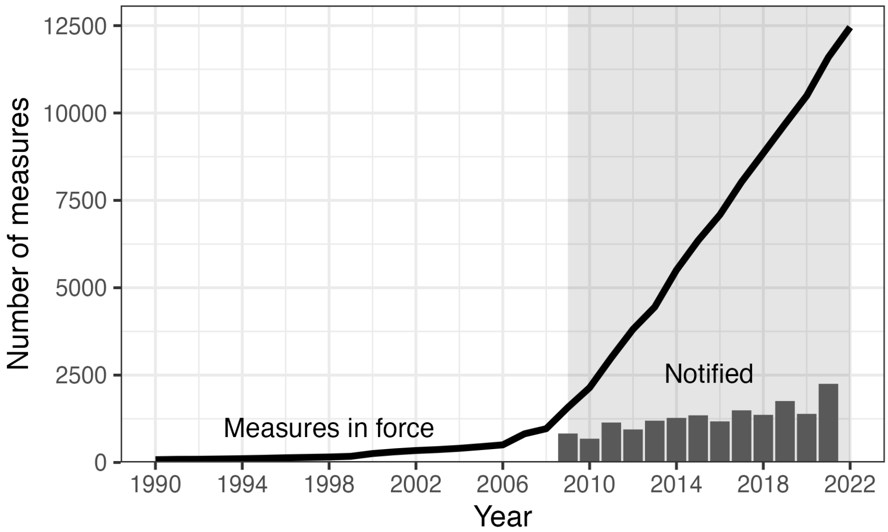
We focused on two key variables in economic policy analysis: 1) the implementation dates of the measures, and 2) the broad sectors of products affected by the measures. Furthermore, given the absence of direct information on the magnitude of policies, we categorized the policy measures and developed a score to capture the strength of a policy measure.

2.2.1 Implementation dates

As it stands, the notification dataset is organized according to the year of notification for each policy measure. However, it is not unusual to receive notifications both before and after the actual implementation date of a policy. Additionally, there can be a time gap between notification and implementation. Given that understanding when a measure comes into effect is often essential for studying the impact of economic policies, our initial objective is to extract this crucial piece of information.

The EDB contains information about the implementation periods of each measure, presented in textual descriptions provided by WTO members in their original notifications. However, these descriptions lack a standardized format, and in many instances, they may encompass multiple

Figure 3: *Number of active policy measures detected in the EDB*



Notes: The figure depicts the increase in active environment-related measures notified to the WTO. The line indicates the number of measures that are in force based on the implementation periods extracted from the EDB. The bar plot illustrate the number of measures by year of notification. We highlighted in grey the notification period covered by the EDB.

implementation periods, expressed in relative terms (e.g., two years after project approval) or even contingent upon other events (e.g., subject to Congressional approval). This heterogeneity complicates the development of an algorithm that can accurately extracts all implementation dates.

We rely on a series of regular expressions to identify starting and ending implementation years by detecting common patterns and phrases in the descriptions. In instances where multiple dates are found within the description, we retain the earliest and latest years as reference for the duration of the measures. To evaluate the accuracy of the algorithm, we conducted a test by randomly selecting 200 measures and manually verifying the extracted dates. In this test sample, only eight years were identified incorrectly. Despite our best efforts, a significant portion of the measures (46%) lack identifiable starting period. This is primarily due to either the absence of implementation dates or insufficient information to establish the date of entry into force, such as with conditional descriptions. In these instances, we assume that the measure entered into force in the year of notification.

The result of our work is visualized in Figure 3, which depicts the number of EDB measures in force and compares it with the number of notifications received each year. It shows that the number of environment-related measures has steadily increased over time, and that by 2009 — the first notification year — about 1400 measures were already in force. For readers interested in further details, we provide in Appendix A the main steps and regular expressions used in our analysis.

2.2.2 Identifying affected sectors

The second part of our extension involves identifying the HS2 sectors affected by the measures. Due to non-uniform reporting obligations across WTO agreements, the structure of notifications and the set of information provided vary depending on the agreement under which the measure was notified. Only about 30% of the measures notified in the EDB report an HS or ICS code describing the goods or standards to which they relate. As shown in Table 1, nearly all of these notifications (74%) are received under the Technical Barriers to Trade Agreement. Having an understanding of the goods affected by the other measures could be useful for economic research.

We aim to extend the sectoral coverage of each measure in the EDB by identifying HS codes for measures for which no product code was provided, and harmonising the sector codes by converting ICS codes to the HS nomenclature.

Table 1: *Number of EDB measures by agreement under which they were notified*

Agreement	**Measures**	**HS/ICS**	%
Technical Barriers to Trade	5343	4033	75.5
Subsidies and Countervailing Measures	4661	31	0.7
Agriculture	3414	3	0.1
Import Licensing Procedures	1644	119	7.2
Quantitative Restrictions	1405	639	45.5
Sanitary and Phytosanitary Measures	1187	604	50.9
Others	552	18	3.3

Notes: For each agreement we report respectively the number of notified measures, the number of notified measures reporting an HS or ICS code, and the share of measures reporting an HS or ICS codes by agreement.

We use natural language processing techniques to parse the description of the measures and identify potential links with HS codes. The linkage is established based on how well the wording in the description matches the products listed in HS chapters. We then use information on the economic sector and environmental goals of the measure to narrow down the potential matches.

Through this approach, we have successfully matched HS 2-digit codes to almost 15,500 measures, bringing the share of measures with an HS codes from 30% to over 85% of all the EDB measures. The remaining measures left unmatched either relate to services (to which HS codes do not apply) or contain only a short or generic description that did not allow our algorithm to match HS codes with sufficient reliability. A full description of our approach is provided in Appendix B, in which we also discuss the conversion of ICS codes to HS and the quality of our final matches.

Table 2 displays the top ten HS chapters affected by EDB measures. Machinery is the category of products to which most of the measures apply. These measures are typically directed to the agricultural sector or aimed at improving energy efficiency and promoting renewable energy. Chemical products are also targeted by a large number of EDB measures. Chemical-related measures are linked to diverse applications, such as agricultural fertilisers, manufacturing activities, packaging, disposal of chemical waste, etc.

Table 2: *Top 10 HS chapters linked to EDB measures*

HS chapter	Freq.	%	Description
84	4441	24.4	Nuclear reactors, boilers, machinery and mechanical appliance; parts thereof
85	3342	18.4	Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers and parts and accessories of such articles
73	1623	8.9	Articles of iron or steel
29	1285	7.1	Organic chemicals
38	1220	6.7	Miscellaneous chemical products.
28	1190	6.5	Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes: Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes
87	1177	6.5	Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof
90	1098	6.0	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof
32	934	5.1	Tanning or dyeing extracts, dyes, pigments, paints and varnishes, putty, and inks.
44	912	5.0	Wood & articles of wood, wood charcoal

2.2.3 Other database extensions

Certain policy measures may be notified multiple times by WTO members, especially in cases where the policy undergoes extension, slight modification, or is simply reconfirmed to be in force. As a result, a single policy could be associated with multiple entries in the EDB dataset, each containing slightly different descriptions and information. To assist in identifying these entries, we calculate a similarity score between all pairs of entries in the dataset. Details are provided in the Appendix C.

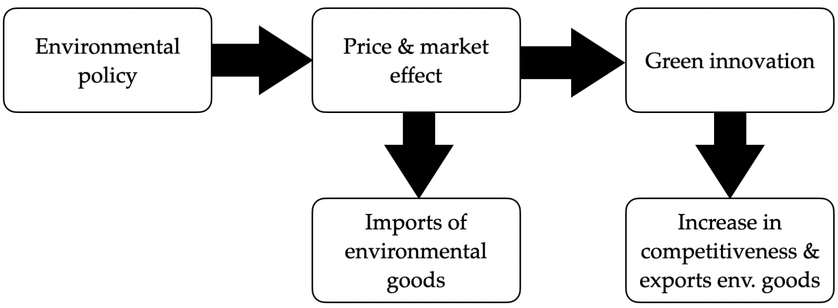
Finally, environmental policies can exhibit significantly diverse designs and stringency. This heterogeneity could pose challenges in producing generalisable results. Therefore, in economic research, it is often beneficial to measure the intensity of a policy to mitigate issues related to unobserved heterogeneity. We endeavor to construct an indicator of policy strength based on the information available in the EDB database. Our measure score is developed along two conceptual dimensions: policy breadth and depth. We define a measure as broad if it impacts a substantial portion of the economy and addresses multiple environmental issues. For our depth scoring, we consider the type of policy instrument used in the measure and the language used in the measure description. It's important to note that quantifying policy measures involves inherent difficulties and subjectivity. Therefore, this policy score should be regarded solely as a proxy for measure strength. A detailed presentation of the scoring system and its calculation is available in Appendix D.

3 Empirical method

The extended dataset enables us to analyse the influence of environmental policies on a range of economic outcomes. In this section, we outline the data and our empirical approach for assessing the effects of environmental policies on innovation and trade in environmental goods.

The potential implications of environmental policies are depicted in Figure 4. We anticipate that environmental policies will stimulate demand for environmental imports. Moreover, these policies are expected to catalyze innovation in environmental technologies, thereby enhancing the competitiveness of the green sector and fostering an increase in the export of environmental goods.

Figure 4: *Potential effects of environmental policy*



3.1 Data

Our analysis centers around three key variables: *innovation*, which we measure with patents data, *trade*, which we measure with the value of merchandise trade, and *environmental policies*, which we derive from the extended EDB dataset. We provide below a brief overview of the data. A description of all the variables and their sources is available in Appendix E.

Environmental innovation

We measure innovation by the number of patents associated with each technology recorded at the IPC subclass level (i.e. 4-digits).[3] Specifically, we employ the fractional count of newly filed patents, drawing on data from OECD (2020). The fractional count method considers the proportional contributions of each country when a patent is jointly invented by individuals from multiple countries. We take the priority date (date of application in the first patent office) as date of reference and consider that innovation took place at the inventor's country of residence. Patent data is commonly used in innovation studies and is considered a better proxy for innovation than other economic variables such as R&D expenditure or the number of active researchers (OECD, 2009).[4]

We focus on patents in the "triadic family" and "IP5 family". The former is a set of patents filed at the European Patent Office (EPO), the Japan Patent Office (JPO) and the United States Patent and Trademark Office (USPTO). The latter is a set of patents filed in at least two patent offices among the five largest IP offices, i.e., EPO, JPO, USPTO, the Korean Intellectual Property Office (KIPO), and the National Intellectual Property Administration of People's Republic of China (NIPA). Restricting patents to these two families helps mitigate the home-bias issue and selects patents of higher quality because only valuable innovations are considered worth the higher cost of patenting across multiple foreign jurisdictions (OECD, 2009). Therefore, patents within the triadic and IP5 families capture innovations considered significant and with international implications.

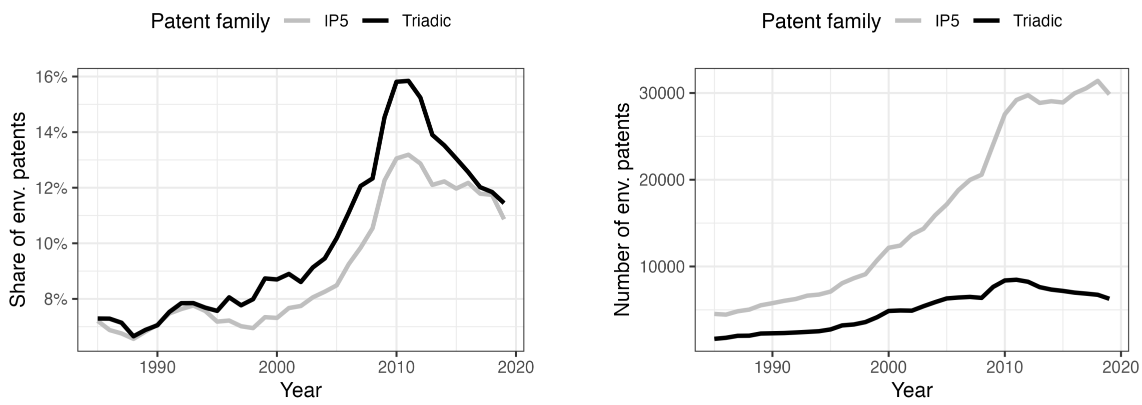
As depicted in Figure 5, the proportion of environmental patents within the triadic family and IP5 family exhibited a gradual increase over time, peaking at about 13-16% around 2011, followed by a subsequent decline to around 11% in 2019. Regarding the total number of patents, environmental patents in the IP5 family surpass those in the triadic family, owing to the relatively less stringent definitions applied to patents in the IP5 family.

To consolidate a unified dataset for our analysis, we use the HS-IPC concordance table developed by Lybbert & Zolas (2014) to link the patents data (4-digits IPC codes) with the HS 6-digit codes from the trade data. The tables of Lybbert & Zolas (2014) also provide the probability of linkage between IPC and HS, which we use to adjust the number of patents in the concordance process.

[3] The International Patents Classification (IPC) is a system used to categorise patents. More information on the IPC can be found at: https://www.wipo.int/classifications/ipc/en

[4] Nonetheless, it should be noted that patents do not capture certain types of innovation, such as learning by doing and informal innovation, that could be taking place in relation to environmental transition (Dechezleprêtre & Glachant, 2014).

Figure 5: *Share and number of environmental patents*



(a) *Share of environmental patents* ***(b)*** *Number of environmental patents*

Notes: This figure shows the proportion of environmental patents within the triadic family and IP5 family (left) and the total number of environmental patents (right). Environmental technologies defined by Haščič & Migotto (2015).

Trade in environmental goods

We measure the value of merchandise trade at the HS 6-digits level using data from the International Trade Database at the Product-Level (BACI) compiled by Gaulier & Zignago (2010). To identify environmental products, we reference the list of environmental goods in Sauvage (2014), which comprises 247 HS 6-digit codes aligned with environmental policy objectives such as air pollution control, water management, environmental monitoring, or the adoption of renewable energies.

Since the environmental products are classified using the 2007 version of the Harmonized Commodity Description and Coding System, commonly known as the Harmonized System (HS), we align the trade data in BACI to correspond with the 2007 version of HS codes. For years predating 2007, this harmonization process involves utilizing a concordance table from the earlier nomenclature (HS 1992) to align with HS 2007. Following this alignment, we aggregate all environmental goods and relevant innovation identified at the HS 6-digit level to establish an "environmental sector" within each HS 2-digit sectors. This data is subsequently matched with environmental measures associated with each HS2 sector in the EDB.

As our focus lies in gauging the relative competitiveness of environmental exports, we also construct the revealed comparative advantage (RCA) index for exports in environmental goods. The environmental RCA index is derived as the relative share of a country's exports in sector k compared to the relative world share of exports in the same sector:

$$RCA_{ikt} = \frac{X_{ikt}/X_{it}}{X_{kt}/X_t}$$

In the formula above, X_{ikt} in the numerator indicates a country's exports in a sector k in year t, X_{it} indicates the country's total exports in a year. In the denominator, X_{kt} indicates the total world export of goods in sector k in year t and X_t indicates total world exports in year t.[5]

[5]Subsequently, as we compute the average treatment effect of the environmental policy, the RCA index is calcu-

It's important to acknowledge that many environmental products may serve dual purposes, meaning they can be used for both environmental and potentially polluting activities. For instance, car parts can be utilized in both low-emission and high-emission engines. Effectively addressing this challenge requires access to trade data with finer granularity, ideally at the 8-digit or 10-digit level of the HS classification. Unfortunately, such detailed data is not uniformly accessible across all countries. As a result, our analysis may inadvertently exaggerate the extent of trade in environmental goods and the impact of policy measures aimed at fostering their growth.

Environmental policy

The environmental policy measures are derived from the extended EDB, which contains a broad spectrum of policy instruments. To facilitate our analysis, we categorize the various policy measures into two primary groups based on their expected impact on compliance costs. The first group primarily comprises environmental regulations, standards and taxes (henceforth REG measures). The second group primarily consists of various forms of monetary and non-monetary subsidies (henceforth SUB measures). The frequency and the classification of policy instruments into these two groups is shown in Figure 6.

3.2 Empirical strategy

Our empirical strategy needs to address several empirical challenges. First, the implementation of environmental policies is not exogenous due to potential biases arising from unobserved characteristics and/or reverse causality. Second, the impact of environmental policies may vary over time. We use a synthetic counterfactual method as proposed by Arkhangelsky et al. (2021) to address these challenges.

We consider a treatment to be the first time a country implements a policy recorded in the EDB. In general terms, we define the effect of policy j in a country i, environmental sector k, in period t as follows:

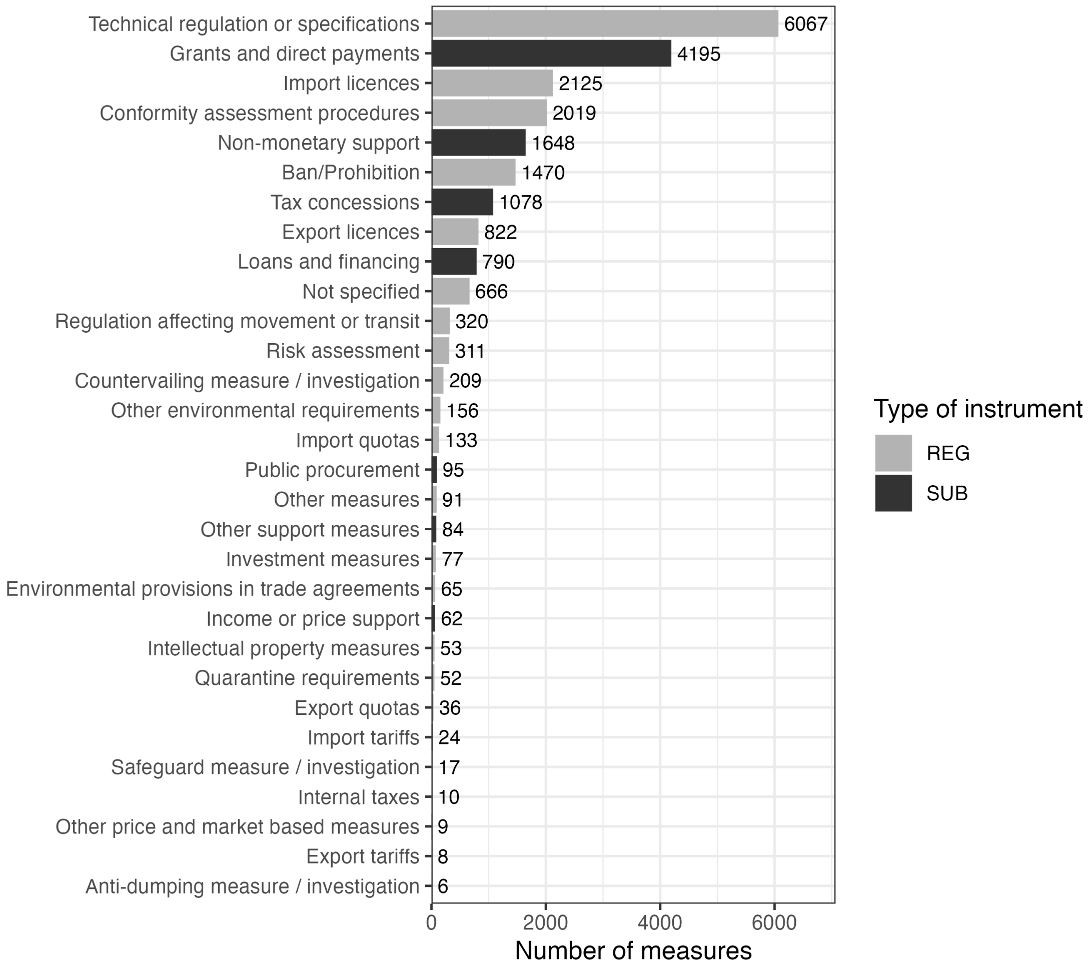
$$\mathrm{TT}_{ijkt} = 100 \cdot \left(\frac{Y_{ikt} - Y^{*}_{ijkt}}{Y^{*}_{ijk,t=T_j}} \right), \quad \forall ik \in \mathrm{Tr}_j \tag{1}$$

In this context, Tr_j is the set of treated country-sectors for policy j, which is composed of all the countries in which the policy is implemented and *environmental* HS2 codes that are linked to the policy according to our matching (Appendix B). $Y^{\star}$ denotes the value for the counterfactual, that is to say, the value of Y if the policy were not enacted. $Y^{*}_{ik,t=T_j}$ designates the value of the counterfactual at the time the policy is enacted (T_j). We use this term to scale the policy effect so that the treatment effects are expressed in percentage for ease of comparison across variables.

Defining the counterfactual is where the challenge lies. Since we cannot directly observe Y^{*}, it is common in the economic literature to rely on the observed Y in a group of untreated country-sectors (*control group*) to formulate a counterfactual and infer the effect of the policy on the treated unit (treatment effect on treated unit, TT). Common strategies involve using the most similar

lated for each country within each environmental HS2 sector. Thus, the RCA index serves to capture competitiveness within environmental sectors.

Figure 6: *Frequency of instruments used in REG and SUB measures*



control unit (e.g. matching strategies), or the evolution in the average of the control group (e.g. difference-in-difference approach).

Ultimately, the validity of these methods hinges on the "parallel trends" assumption, which suggests that trade and patents in the counterfactual would follow the same trajectory if the environmental policies were not enacted. In selecting the control group, we have several options: 1) comparing the outcome variables for the same sectors in different countries, 2) comparing the outcome variables for a different sector in the same country, or 3) comparing the outcomes for different country-sector units. We believe that given the similarity of trade and innovation patterns in the same sectors across countries, the first comparison is the best approach to capture the impact of environmental policy measures. Nonetheless, the parallel trends assumption might not be exactly met for all treated units, therefore, we use a synthetic counterfactual method to match pre-treatment trends between our treated units and the synthetic control, thereby establishing a more credible counterfactual.

For each country affected by policy j in an environmental sector k, we construct a synthetic counterfactual $\hat{Y}^*$ by assuming that, in the absence of a policy, the evolution of Y in the treated unit is well approximated by a linear combination of the observed Y in all the control units. In other words, we select a synthetic control that makes the pre-treatment trends for treatment and control groups as parallel as possible. This approach allows us to circumvent the challenges on two-way fixed effects with differential treatment timing raised by recent work (Callaway & Sant'Anna, 2021).

$$\hat{Y}^{*}_{ijkt} = \omega^{(0)}_{ijk} + \sum_{c \in C_k} \omega^{(c)}_{ijk} Y_{ckt} \tag{2}$$

In the equation above, C_k is the set of control units for policies affecting sector k. It is composed exclusively of countries for which the sector k never experienced a policy recorded in the EDB. $\omega^{(c)}_{ijk}$ are the weights assigned to each control unit c for building the counterfactual of policy j for the treated country-sector ik. Y_{ckt} is the value observed in control country c for sector k at time t. Therefore, the counterfactual is only built from values observed for other countries in the same sector. Finally, for each policy j, $\omega^{(0)}_{ijk}$ is an intercept specific to each treated country-sector, which allows the counterfactual to be parallel instead of requiring an overlap (Abadie & Gardeazabal, 2003; Alberto Abadie & Hainmueller, 2010).

The counterfactual construction follows the methodology developed by Arkhangelsky et al. (2021). The weight ω and $\omega^{(0)}$ are derived through the following constrained optimization problem:

$$\begin{aligned}
\left(\hat{\omega}^{(0)}_{ijk}, \hat{\omega}_{ijk}\right) &= \underset{\omega^{(0)}_{ijk} \in \mathbb{R}, \omega_{ijk} \in \Omega_k}{\arg\min} \ \ell_{\text{unit}} \left(\omega^{(0)}_{ijk}, \omega_{ijk}\right) \text{ where} \\
\ell_{\text{unit}} \left(\omega^{(0)}_{ijk}, \omega_{ijk}\right) &= \sum_{t=1}^{T_{pre}} \left(Y_{ikt} - \left(\omega^{(0)} + \sum_{c \in C_k} \omega^{(c)}_{ijk} Y_{ckt} \right) \right)^2 + \zeta^2_{jk}(T_{pre}) \|\omega_{ijk}\|^2_2. \\
\Omega_k &= \left\{ \omega_{ijk} \in \mathbb{R}^{|C_k|}_{+} : \sum_{c \in C_k} \omega^{(c)}_{ijk} = 1 \right\}.
\end{aligned} \tag{3}$$

where $\zeta^2_{jk} T_{pre} \|\omega_{ijk}\|^2_2$ represents a regularization term common in the synthetic counterfactual literature. We follow Arkhangelsky et al. (2021) in setting the penalty coefficient (ζ) to be equal to the size of a typical one-period outcome change Δ_{ijkt} for unexposed units in their pre-treatment period, multiplied by a theoretically motivated scaling $(|C_k| \cdot (T_{post}))^{1/4}$:

$$\begin{aligned}
&\zeta_{jk} = (|\text{Tr}_j| \cdot (T_{post}))^{1/4} \, \hat{\sigma}_{jk} \text{ with } \hat{\sigma}^2_{jk} = \frac{1}{|C_k| \cdot (T_{pre} - 1)} \sum_{c \in C_k} \sum_{t=1}^{T_{pre}-1} \left(\Delta_{ckjt} - \bar{\Delta}_{jk}\right)^2, \\
&\text{where} \quad \Delta_{cjkt} = Y_{cjk(t+1)} - Y_{cjkt}, \quad \text{and} \quad \bar{\Delta}_{jk} = \frac{1}{|C_k| \cdot (T_{pre} - 1)} \sum_{c \in C_k} \sum_{t=1}^{T_{pre}-1} \Delta_{cjkt}.
\end{aligned} \tag{4}$$

In the equation above, T is the total number of time periods in our data sample, $T_{pre} = T_j - 1$ is the number of pre-treatment periods, and $T_{post} = T - T_{pre}$ is number of post-treatment periods for policy j. Finally, $|\text{Tr}_j|$ and $|C_k|$ indicate respectively the number of units treated by policy j and the number of control units available for environmental sector k.

We summarise the policy effects by computing the following average treatment effect on the treated (ATT) for different policy lag lengths (l).

$$ATT_l = \frac{1}{J} \sum_{j=1}^{J} \frac{1}{|\text{Tr}^{(i)}_j|} \sum_{i \in \text{Tr}^{(i)}_j} \sum_{k \in \text{Tr}^{(k)}_j} \tilde{L}_{jk} \cdot \text{TT}_{ijk,t=l-T_j}, \tag{5}$$

Where J is the total number of policies in the EDB, $\text{Tr}_j^{(i)}$ and $\text{Tr}_j^{(k)}$ are respectively the set of countries and sectors that are affected by policy j. The equation corresponds to a weighted average of all the treatment effects computed for each policy-country-sector. Each policy is assigned an equal weight and within each policy, every country is given equal weighting. Conversely, the contribution of each sector is weighted by the relative link strength ($\tilde{L}_{jk}$) between the policy and the sector described in detail in Appendix B. This metric captures the likelihood that the measure relates to a given HS chapter.

An additional detail in the implementation is worth nothing. Given the considerable variation in size among countries, the magnitude in the trade and patent data varies significantly across countries. This can significantly influence the quality of the counterfactual because we constrain the weights ω to add up to one and because the regularization term depends on the variance across units (through σ). Therefore, we follow the recommendation of Abadie (2021) to rescale the variables to a common unit prior to computing the counterfactual. We standardized all variables by computing their z-score.[6] The results are then presented by converting the z-score into the original value of the outcome variables.

3.3 Results

The results of synthetic counterfactual estimation on trade patterns are presented in Figure 7. The horizontal axis shows the number of years around the implementation of the measures, with 0 indicating the year the measure was introduced. The vertical axis denotes the impact of the environmental measure on the treated country-sector relative to the synthetic counterfactual. The results represent the average treatment effects on the treated (ATT) in equation 5. The value is expressed in percentage change relative to the observed outcome in the year of the policy implementation.

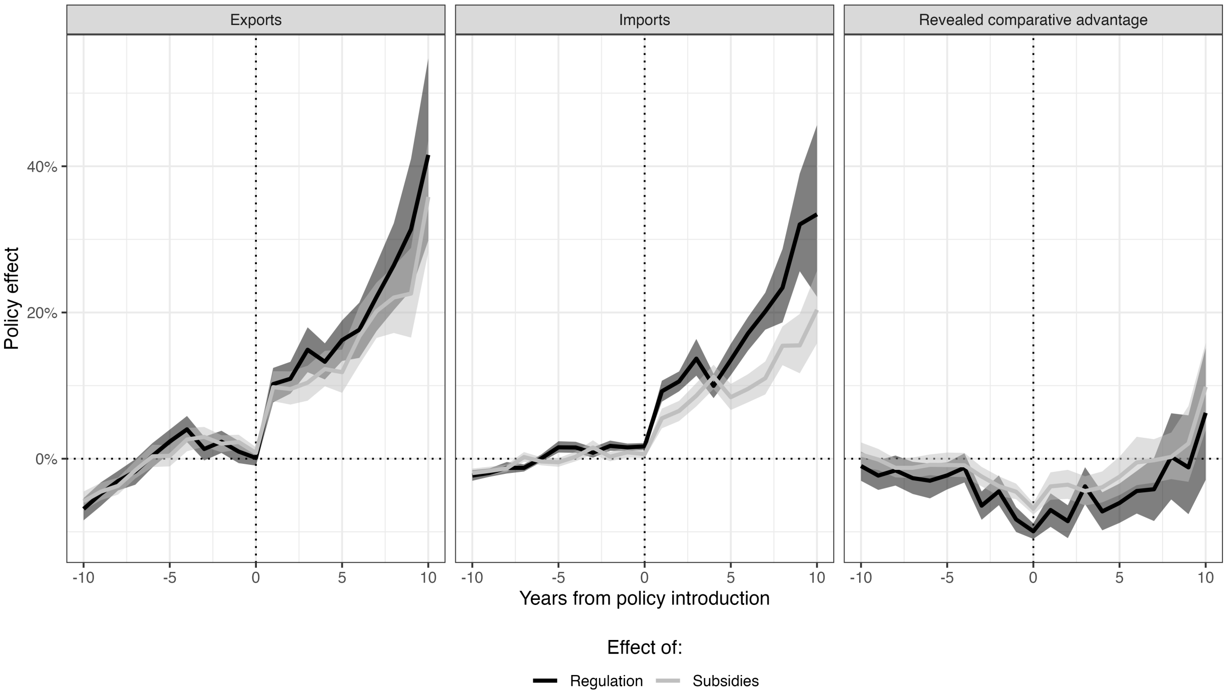
Both the REG and SUB policies demonstrate significant impacts on a country's exports and imports. The effects gradually increase over time after the policy's implementation, and are associated with an increase in exports of up to 40% and a 30% increase in imports of environmental goods relative to the synthetic counterfactual approximately ten years after the environmental measures are enacted. Moreover, REG and SUB policies display subtle differences in their trade effects: while both REG and SUB policies appear to have strong effects on the exports of environmental goods, REG measures tend to boost imports slightly more.

The increase in the trade of environmental goods, however, may be due to a general increase in trade values. When measured by the revealed comparative advantage (RCA) of environmental exports, the effect do not appear to be as pronounced. The implementation of an environmental policy does not appear to be associated with an increase of the comparative advantage of a country's exports in the initial years of a policy's implementation, and the effect is only visible eight to ten years after a policy is enacted.

We also present some illustrative evidence on the policies' impacts on green innovation in Figure 8, which comprises patents in both the triadic and IP5 families. A notable and statistically significant increase in innovation is observed in the years following the implementation of an environmental policy reflected in the number of patents in the triadic family and the IP5 family.

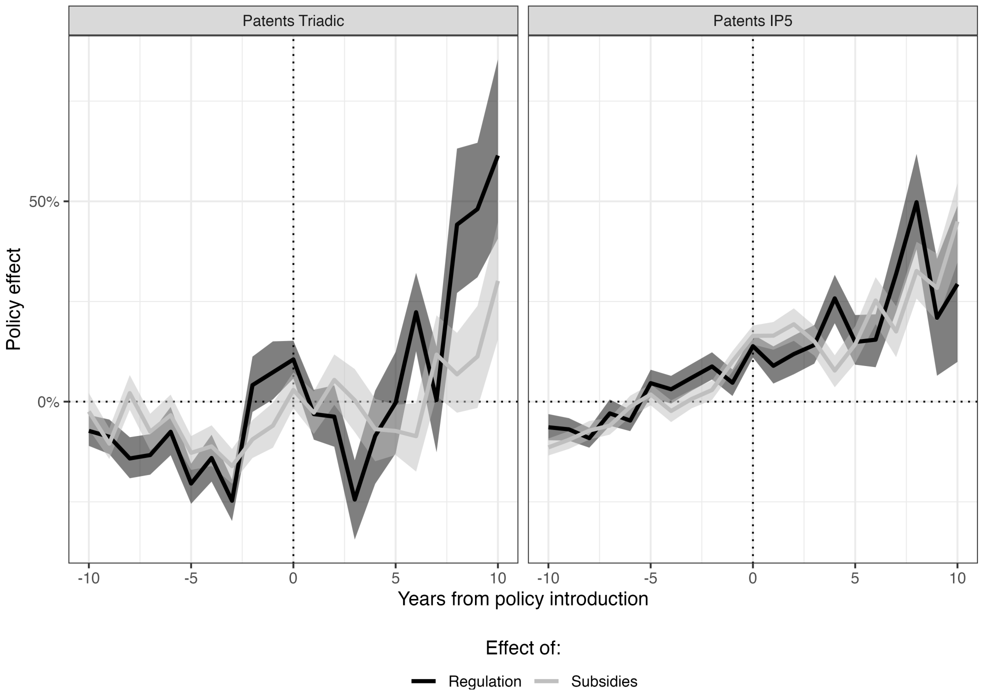
[6] The z-score is calculated as $(Y_{ikt} - \mu(Y)_{ik})/\sigma(Y)_{ik}$, where $\mu(Y)_{ik}$ is the mean of the variables, and $\sigma(Y)_{ik}$ is its standard deviation for the country-sector ik.

Figure 7: *Potential effects of environmental policy on trade*



Notes: The figure displays the estimation results of synthetic counterfactual to illustrate the effects of environmental policies on (1) exports of environmental goods, (2) imports of environmental goods and, (3) revealed comparative advantage in environmental goods exports. The horizontal axis represents the years around the implementation of an environmental policy, while the vertical axis indicates the percentage change in the trade in the country implementing an environmental policy compared with trade in the same sector in the counterfactual countries with no policy. The shaded area represents bootstrapped 95% confidence interval.

Figure 8: *Potential effects of environmental policy on green innovation*



Notes: The figure displays the estimation results of synthetic counterfactual to illustrate the effects of environmental policies on environmental patents within the triadic family and IP5 family. The horizontal axis represents the years around the implementation of an environmental policy, while the vertical axis indicates the percentage change in patents in the country implementing an environmental policy compared with patents in the same sector in the counterfactual countries with no policy. The shaded area represents bootstrapped 95% confidence interval.

While the patent counts in the triadic family seems to display some fluctuations, the patent counts in the IP5 family appears to show a steady increase after the implementation of an environmental policy. While both REG and SUB policies appear to have positive impacts on the filing of patents, SUB measures seem to have a stronger impact on green innovation compared with REG measures.

These average effects mask the heterogeneous impacts of the policy measures by country, sector, and policy instruments. Appendix F presents the potential effects of environmental policies on revealed comparative advantage of environmental goods exports by country. Furthermore, we present additional estimated impacts of each individual type of environmental measure in Appendix F.

In summary, we employ a synthetic counterfactual method to address the empirical challenges in our dataset and shed light on the potential impacts of environmental policies. Our empirical analysis suggests that the environmental measures recorded in the EDB have a positive and significant impact on both the exports and imports of environmental goods. Additionally, these measures appear to stimulate green innovation.

While we have made efforts to address endogeneity issues with our empirical strategy, it is crucial to acknowledge the method's limitations. First, we recognize that a multitude of other factors can simultaneously affect a country's decision to implement an environmental policy as well as its trade and innovation patterns. While our method can capture these unobserved effects if the trends are similar between countries in the treated and control groups, it does not address the issue if these factors exclusively affect only a subset of the countries in the post-treatment period. Second, the effects of environmental measures are not limited to one sector or one country and may spill over to other countries. We believe that by comparing the effects of the treatment on the same HS2 sectors in different countries, our method best addresses the potential spillover effect across sectors. Yet, if the policies' effects spill over to countries in the control group, this can result in a violation of the stable unit treatment value assumption crucial in such analysis. Third, since we estimate policy effects only from notified measures in the EDB, our estimates could be biased if there is systemic under-reporting of certain types of policy measures. Due to these considerations in our empirical method, we emphasize that the results should be interpreted as indicative of correlation rather than causation.

4 Conclusion

This study contributes to the economic literature in two key ways. Firstly, we utilized text analysis algorithms to extract valuable information from the WTO environmental database (EDB), enhancing the dataset's utility for future research and policy analysis. The methodological innovation in discerning the implementation year and applicable sectors of policy measures may also hold significant value for future research endeavors in examining the impact of policy measures.

Secondly, we employed this extended dataset to investigate how environmental measures affect environmental trade and green innovation. Our findings indicate that environmental policies not only stimulate imports of environmental goods but also enhance competitiveness and boost exports of environmental goods. Furthermore, our results suggest that environmental policies can foster innovation. Thus, our findings underscore the dynamic impacts of environmental policies on trade and innovation, emphasizing the importance of policy design.

Given the increasing trend of countries enacting environmental policies, understanding the environmental and economic impacts of these measures is of crucial importance. Our efforts to utilize a public dataset to estimate the impact of environmental measures represent a step in this direction. Further work to examine different policy instruments and their interactions holds promise for guiding policymakers in crafting more effective policies for sustainable development.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2), 391–425.

Abadie, A. & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1), 113–132.

Acemoglu, D., Aghion, P., Bursztyn, L., & Hemous, D. (2012). The environment and directed technical change. *American Economic Review*, 102(1), 131–166.

Acemoglu, D., Aghion, P., & Hémous, D. (2014). The environment and directed technical change in a North–South model. *Oxford Review of Economic Policy*, 30(3), 513–530.

Acemoglu, D., Akcigit, U., Hanley, D., & Kerr, W. (2016). Transition to clean technology. *Journal of Political Economy*, 124(1), 52–104.

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., & Van Reenen, J. (2012). Carbon taxes, path dependency and directed technical change: Evidence from the auto industry. *NBER Working Paper*, No. 18596.

Alberto Abadie, A. D. & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505.

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.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12), 4088–4118.

Babiker, M. H. (2005). Climate change policy, market structure, and carbon leakage. *Journal of International Economics*, 65(2), 421–445.

Bettarelli, L., Furceri, D., Pizzuto, P., & Shakoor, N. (2023). *Environmental Policies and Innovation in Renewable Energy*. Technical report, International Monetary Fund.

Bretschger, L. & Suphaphiphat, N. (2014). Effective climate policies in a dynamic North–South model. *European Economic Review*, 69, 59–77.

Burghaus, K. & Funk, P. (2013). Endogenous growth, green innovation and GDP deceleration in a world with polluting production inputs. In *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Growth and the Environment*, C04-V3: ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft. Available from http://hdl.handle.net/10419/80022.

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

Callaway, B. & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2), 200–230.

Copeland, B. R. & Taylor, M. S. (2004). Trade, growth, and the environment. *Journal of Economic Literature*, 42(1), 7–71.

Dechezleprêtre, A. & Glachant, M. (2014). Does foreign environmental policy influence domestic innovation? Evidence from the wind industry. *Environmental Resource Economics*, 58, 391–413.

Dechezleprêtre, A. & Sato, M. (2017). The impacts of environmental regulations on competitiveness. *Review of Environmental Economics and Policy*, 11(2), 183–206.

Di Maria, C. & Smulders, S. (2004). Trade pessimists vs technology optimists: Induced technical changeand pollution havens. *Advances in Economic Analysis & Policy*, 4(2).

Duan, Y., Ji, T., & Yu, T. (2021). Reassessing pollution haven effect in global value chains. *Journal of Cleaner Production*, 284, 124705.

Eugster, J. (2021). *The impact of environmental policy on innovation in clean technologies*. Technical report, International Monetary Fund.

European Commission (2021). Legal sources on renewable energy (RES LEGAL). *European Commission*. Accessed in February 2021. Available from http://www.res-legal.eu/.

Gaulier, G. & Zignago, S. (2010). *BACI: International Trade Database at the Product-Level. The 1994-2007 Version*. Working Papers 2010-23, CEPII.

Greaker, M., Heggedal, T.-R., & Rosendahl, K. E. (2018). Environmental policy and the direction of technical change. *The Scandinavian Journal of Economics*, 120(4), 1100–1138.

Han, X., Schmidt, J., & Steingress, W. (2019). Constructing a Concordance Table between Harmonized System (HS) and International Classification for Standards (ICS). *Working paper*. Available from http://julia-schmidt.org/Han_Schmidt_Steingress_ICS_HS.pdf.

Hart, R. (2019). To everything there is a season: Carbon pricing, research subsidies, and the transition to fossil-free energy. *Journal of the Association of Environmental and Resource Economists*, 6(2), 349–389.

Haščič, I. & Migotto, M. (2015). Measuring environmental innovation using patent data. *OECD Environment Working Papers*, No. 89. Available from http://www.oecd.org/env/indicators-modelling-outlooks/green-patents.htm.

Hémous, D. (2016). The dynamic impact of unilateral environmental policies. *Journal of International Economics*, 103, 80–95.

IRENA (2021). Renewable energy target data dashboard. *International Renewable Energy Agency*. Available from https://www.irena.org/statistics.

Kellenberg, D. K. (2009). An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of International Economics*, 78(2), 242–255.

Koźluk, T. & Timiliotis, C. (2016). Do environmental policies affect global value chains? *OECD Economics Department Working papers*, No. 1282. Available from https://www.oecd-ilibrary.org/content/paper/5jm2hh7nf3wd-en.

Levinson, A. & Taylor, M. S. (2008). Unmasking the pollution haven effect. *International Economic Review*, 49(1), 223–254.

LSE/Columbia Law School (2021). Climate Change Laws of the World database. *Grantham Research Institute on Climate Change and the Environment and Sabin Center for Climate Change Law*. Available from https://climate-laws.org/.

Lybbert, T. J. & Zolas, N. J. (2014). Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities' approach for joint analyses of patenting and economic activity. *Research Policy*, 43, 530–542.

Millimet, D. L. & Roy, J. (2016). Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous. *Journal of Applied Econometrics*, 31(4), 652–677.

OECD (2009). OECD patent statistics manual. *Organisation for Economic Co-operation and Development*. Available from http://www.oecd.org/science/inno/oecdpatentstatisticsmanual.htm.

OECD (2020). OECD Environment Statistics database. *Organisation for Economic Co-operation and Development*. Accessed in November 2020. Available from https://www.oecd-ilibrary.org/environment/data/oecd-environment-statistics_env-data-en.

OECD (2020). OECD patent statistics; Patents by main technology and by international patent classification (IPC). *Organisation for Economic Co-operation and Development*. Data accessed in August 2020. Available from https://doi.org/10.1787/data-00508-en.

OECD (2021). Policy instruments for the environment (PINE) database. *Organization for Economic Co-operation and Development*. Accessed in February 2021. Available from https://pinedatabase.oecd.org/.

Peters, M., Schneider, M., Griesshaber, T., & Hoffmann, V. H. (2012). The impact of technology-push and demand-pull policies on technical change — Does the locus of policies matter? *Research Policy*, 41(8), 1296–1308.

Pugliese, E., Cimini, G., Patelli, A., Zaccaria, A., Pietronero, L., & Gabrielli, A. (2019). Unfolding the innovation system for the development of countries: coevolution of science, technology and production. *Scientific reports*, 9(1), 16440.

Sauvage, J. (2014). The stringency of environmental regulations and trade in environmental goods. *OECD Trade and Environment Working Papers*, 2014/03. Available from https://dx.doi.org/10.1787/5jxrjn7xsnmq-en.

Stöckl, F. (2020). Is substitutability the new efficiency? Endogenous investment in the elasticity of substitution between clean and dirty energy. *Deutsches Institut für Wirtschaftsforschung*, Discussion Paper No.1886. Available from http://dx.doi.org/10.2139/ssrn.3671370.

UN (2020). UN Comtrade database. *UN Trade Statistics*. Available from https://comtrade.un.org/.

UNSD (2020). National accounts statistics: Main aggregates and detailed tables, 2018. *United Nations Statistics Division*. Accessed in Febrary 2020. Available from https://data.un.org/.

Witajewski-Baltvilks, J. & Fischer, C. (2018). Green innovation and economic growth in a North–South model. *IBS Working Papers*, 10/2018.

World Bank (2019). World Development Indicators dataset. *The World Bank*. Data accessed in March 2020. Available from https://data.worldbank.org/.

WTO (2020). Environmental database (EDB). *World Trade Organization*. Available from https://edb.wto.org/.

Note: With the exception of Appendix F, all the figures and values reported in the appendices are based on an old version of the EDB dataset containing only measures notified up to 2020. The latest version of the EDB, which is used in the main body of the article, covers measures notified up to 2022.

A Extracting implementation years

The algorithm and regular expressions presented here outline the main steps that we took in extracting the implementation years out of the “Implementation Period” variable of the EDB. Additional data cleaning procedures were also applied to ensure consistency in the extracted dates. Moreover, for some type of measures for which only a starting year is expected (e.g. standards, regulation and taxes), we used a simplified approach that only searched for the starting year.

```
CLEAN TEXT BY KEEPING ONLY DESCRIPTION AFTER:
    ".*((?:Duration of the measure|Duration of the subsidy).*)"

IF NOT FOUND, REMOVE REPORTING DATES BY MATCHING AND KEEPING GROUP 1 OF:
    "^(?:(?:\d{1:2} )?(?:January|February|March|April|May|June|July|August|
    September|October|November|December) )?\d{4} - (?:(?:\d{1:2} )?(?:January|
    February|March|April|May|June|July|August|September|October|November|
    December) )?\d{4}(\[A-z]+.*)"

THEN FIND ANY DATE RANGE IN THE TEXT BY MATCHING:
    "(?:[Ff]rom|[Ss]ince|[Bb]etween)?(?:(?: \d{1,2}[stndrh]{0,2})?\s?(?:of)?
    [a-zA-Z]{0,9})?\s?(\d{4}) (?:to|until|up to|-|till|and)(?:(?: \d{1,2}
    [stndrh]{0,2})?\s?(?:of)?[a-zA-Z]{0,9})? (\d{4})\b"

IF ANY DATE RANGE WAS FOUND, KEEP THE LOWEST AND HIGHEST YEAR IDENTIFIED.

LOOK FOR THE PRESENCE OF MEASURE END DATES:
    "(?:Ends|[Ee]nded(?: on| in)?|[Ee]nding(?: on|in)?|[Ee]xpire[sd](?: on|
    in)?|Terminated(?: on| in)?|available until|Until|[Uu]ntil end|Prolonged
    until|[Uu]p to the end(?: of)?|Project completed after|On-going until|
    until and including|Currently to|will expire on|not be applied after the
    year|available till|repealed\s?for facilities placed in service after|
    continue provisionally until|Phase-out from|produced before|On-going
    [[:punct:]] sunset|[Ss]unset[[:punct:]]|Last date for application
    [[:punct:]]|[Ss]unsets(?: in| on)?|[Ee]xpiration of the [Ll]aw(?: on| in))
    (?:(?:\d{1,2}[stndrh]{0,2})?\s?(?:of )?[a-zA-Z]{0,9}\s?)?(?:\d{1,2}/
    \d{1,2}/)?(\d{4})\b"

CHECK FOR SINGLE YEAR MEASURES:
    "(?:(?:Calendar|Fiscal|Marketing|Financial) year|FY) (?:(?:\d{1,2}[stndrh]
    {0,2})?\s?(?:of )?[a-zA-Z]{0,9}\s?)?(?:\d{1,2}/\d{1,2}/)?
    (\d{4})$|^(?:[Dd]uration of the (?:subsidy|measure|policy):(?: [Tt]he
```

```
    [Yy]ear)?\s)?(\d{4})$"

LOOK FOR SMALLEST YEAR TO USE AS START YEAR IF NONE WAS PREVIOUSLY FOUND:
    "(?:[[:punct:]]|\b)(\d{4})(?:[[:punct:]]|\b|Period of application|
    Duration of the)"

LOOK FOR LARGEST YEAR TO USE AS END YEAR IF NONE WAS PREVIOUSLY FOUND:
    "(?:\b|[[:punct:]])(\d{4})(?:\b|[[:punct:]])"

IF NO TEXT INFORMATION WAS PROVIDED, NO DATE WAS MATCHED, OR IF THE MEASURE ONLY
HAS THE END DATE, USE THE NOTIFICATION YEAR AS STARTING YEAR

IF NO ENDING YEAR WAS IDENTIFIED, ASSUME IT HAS INDEFINITE APPLICATION
```

Figure 9 illustrates the number of EDB measures in force and compares it with the number of notifications received each year. It shows that the number of environment-related measures has steadily increased over time, and that by 2009 — the first notification year — about 1400 measures were already in force.

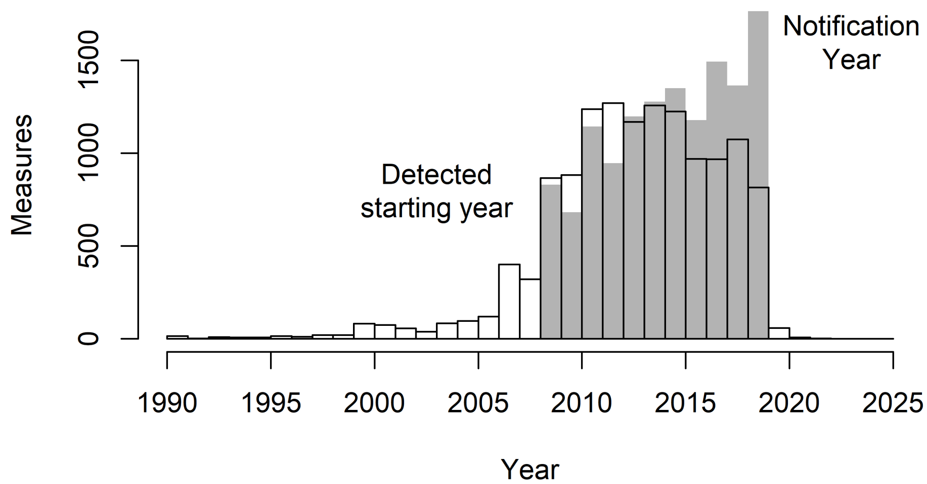


Figure 9: *Comparison of the notification and detected starting years*

Notes: The plot displays the number of measures by notification (shaded bars) and detected starting year (empty bars).

B Linking HS codes to EDB measures

In this note we describe how we matched HS codes to the measures in the EDB database. The goal of this method is to use the information included in the text description of the variables "coverage of measure", "measure description" and "environment related objective" to find possible matches for the measures. This methodology closely follows the one of Han et al. (2019), with a few additions to incorporate information from multiple sources and adapt it to our matching problem.

The basic idea consists in calculating a score that represents the likelihood of environmental measures being linked to a specific HS code. This score, which we call link strength, is calculated from the number (and specificity) of keywords that are found in the description of both the measure and the HS category. This score is then adjusted to take into account how likely the HS code is to be linked to the harmonised economic sector and environmental objective of the measure. Eventually, only the strongest links are kept.

Step 1: Extracting and cleaning keywords

We start by extracting every single word out of the description of the HS categories and EDB's combined three columns: "measure description", "coverage of the measure" and "Environment related objective". These words are then reduced to their root form (e.g. wood for wooden). To do this, Han et al. (2019) uses a stemming algorithm, but we opted for a lemmatisation algorithm. Stemming is faster, since it works by truncating words, but lemmatisation usually produces a better result because it refers to a dictionary to find the root form of words. We use the `udipipe`[7] package in R to perform the tokenisation and lemmatisation of the descriptions. This package also allows us to annotate useful information about the part-of-speech categories (e.g. verbs, nouns, adverbs, etc.) of each word, as well as it's role within each sentence (e.g. clausal subject, object, etc.).

To simplify the list of keywords and keep only the most informative, we decided to keep exclusively keywords that are flagged as nouns, verbs, adjectives or proper nouns. We also ensured everything is in lowercase and removed all stop words. Stop words are common words in a language that usually do not carry substantial information (e.g. the, a, in). We use the Snowball list[8] as a base and expand it with generic policy words that we have found to be particularly influential during the matching. The complete list of words we manually added is found in Table 3.

Step 2: Linking measures and HS categories

For every notified measure in the EDB (i), a link is established with the HS 2-digits categories (j) which shares at least one keyword in common. From now on, the keywords of the HS classification are grouped at the 2-digits level. That is to say, the keywords extracted from the HS 6-digits, 4-digits and 2-digits description are all grouped together to describe the HS chapter. Let N_{ik} be the frequency of a keyword k in description of the measure i and in the same fashion N_{jk} the frequency of keywords in the HS category j. Then, the strength of the link L is measured by:

$$L_{ij} = \sum_{k=1}^{K_i} N_{ik} \cdot (N_{jk} \cdot \omega_k)$$

[7]The package is available from `https://cran.r-project.org/web/packages/udpipe/index.html`.
[8]The list of words is available from `http://snowball.tartarus.org/dist/snowball_all.tgz`.

Table 3: *Policy stop words*

act	condition	implement	number	protect	state
active	control	include	objective	protection	support
address	country	individual	operation	provide	system
aid	current	intend	order	public	technical
apply	define	issue	particular	reduce	trade
area	develop	large	payment	register	value
basic	development	level	person	regulate	year
better	draft	low	plan	result	
business	facility	maximum	producer	small	
certain	framework	medium	programme	specific	
chapter	group	method	project	specify	
commercial	high	new	property	standard	

The expression above describes how the strength of the link (L) is calculated by summing for every distinct keyword k, out of the K_i total number of distinct keywords in the description of the measure i, the product of the frequency of the keyword in the description of i and j. The product of the two frequencies will associate higher scores whenever the keyword appears multiple times, reflecting the fact that they are more important in the description.

As in Han et al. (2019), a TF-IDF[9] weighting scheme is introduced to highlight the most important words for the specific HS 2-digit category. This weighting (ω) gives more importance to words which are specific to single HS chapter. It is defined for the keyword k in the following way:

$$\omega_k = 1 + \log\left(\frac{1 + J^*}{1 + J_k}\right)$$

Where J^* is the total number of HS 2-digits categories and J_k is the number of HS categories which contain the keyword k. Given that in our data there are 97 distinct HS categories J^*, the weight ω ranges between 1 and approximately 4.9.

Finally, we also apply a weighting factor to specific keyword-chapter combinations which we found to be dominant in the data sample and not particularly representative of the HS chapter content (Table 4).

Step 3: Incorporate information from the harmonised sectors and objectives

At this stage, we obtained all possible HS categories to which the measures are linked and calculated the strength of this linkage L. Now the information provided in the variable "harmonised sector" and "harmonised environmental objective" can be used to eliminate less relevant links and increase the precision of the matching.

The variable "harmonised sector" contains a description of the broad economic sectors that are affected by the measure i (e.g. agriculture, fisheries, chemicals, energy, manufacturing, mining, etc.). These harmonised sectors could be matched to the HS chapters in the way described in Table 5. This table establishes a rough correspondence between HS chapters and sectors of economic activity. We use it to help identifying the most likely links among the ones we found in step 2.

[9]Term Frequency - Inverse Document Frequency (TF-IDF)

Table 4: *Reducing sensitivity to influential words in certain chapters*

Word	**Chapters**	**Weight**
water	84	0.2
	3, 69, 7	0.3
gas	7, 84, 85	0.3
air	84	0.3
special	87	0.5
design	87	0.5
agricultural	87	0.5
oil	85	0.3
plant	84	0.3
production	90	0.3
safety	70	0.5
consumption	3	0.5

Table 5: *Tentative matching of Harmonised sectors and HS chapters*

Harmonised sector	HS chapters
Specific sectors:	
Agriculture	6–14
Chemicals	28–40
Energy	84–85
Forestry	44–48
Fisheries	3
Manufacturing	15–24, 50–70, 84–96
Mining	25–27, 71–83
Other sectors:	
All products/economic activities	1–97
Not specified	1–97
Other	1–2, 4–5, 41–43, 49, 97–99
Services	—

In a similar fashion, the variable "harmonised environmental objectives" provides useful information on the type of environmental objective that is targeted by the measure. This information can be combined with the OECD list of environmental goods (Sauvage, 2014) to narrow down the HS codes related to the measure. The OECD list of environmental goods records a series of goods (and their respective HS codes) that are used to achieve specific environmental goals, such as air pollution control, waste management or animal protection. Again, a correspondence is established between the "harmonised environmental objectives" of the EDB database and the environmental goals of the OECD list. The full correspondence table is presented in Table 6.

Table 6: *Environmental objectives and OECD's environmental goods*

OECD category	OECD product type	Associated HS chapters	Harmonised environmental objective
Air pollution control	Air-handling equipment	84	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; General environmental protection
	Catalytic converters	84	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management
	Chemical recovery systems	25, 28, 84, 38	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; General environmental protection
	Dust collectors	84	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management
	Incinerators, scrubbers	84, 85	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management
	Odour control equipment	84	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management
	Separators/precipitators	70, 84	Air pollution reduction; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management
Cleaner/resource efficient technologies and processes	Cleaner/resource efficient technologies and processes	28, 32	Air pollution reduction; Climate change mitigation and adaptation; Energy conservation and efficiency; Environmental goods and services promotion; Environmentally friendly consumption; General environmental protection; Natural resources conservation
Environmental monitoring, analysis and assessment	Measuring and monitoring equipment	90	Air pollution reduction; Chemical, toxic and hazardous substances management; Environmental goods and services promotion; General environmental protection

Table 6: Environmental objectives and OECD's environmental goods (continued)

OECD category	OECD product type	Associated HS chapters	Harmonised environmental objective
	Process and control equipment	90	Air pollution reduction; Chemical, toxic and hazardous substances management; Environmental goods and services promotion; General environmental protection
Noise and vibration abatement	Mufflers/silencers	84, 87	Animal protection; Environmentally friendly consumption; General environmental protection
Remediation and cleanup	Cleanup	85, 90	Animal protection; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Environmental protection from pests and diseases; General environmental protection; Plant protection; Soil management and conservation; Waste management and recycling
	Water treatment equipment	85	Animal protection; Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Environmental protection from pests and diseases; General environmental protection; Plant protection; Soil management and conservation; Water management and conservation
Renewable energy plant	Heat/energy savings and management	38, 70, 84, 85, 90	Air pollution reduction; Alternative and renewable energy; Climate change mitigation and adaptation; Energy conservation and efficiency; Environmental goods and services promotion; Environmentally friendly consumption; General environmental protection; Natural resources conservation
	Other	29, 22	Air pollution reduction; Alternative and renewable energy; Climate change mitigation and adaptation; Environmental goods and services promotion; General environmental protection; Natural resources conservation

Table 6: Environmental objectives and OECD's environmental goods (continued)

OECD category	OECD product type	Associated HS chapters	Harmonised environmental objective
	Solar	84, 85	Air pollution reduction; Alternative and renewable energy; Climate change mitigation and adaptation; Environmental goods and services promotion; General environmental protection; Natural resources conservation
Solid waste management	Hazardous waste storage and treatment equipment	68, 78, 85, 90	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; General environmental protection; Plant protection; Soil management and conservation; Waste management and recycling
	Waste collection equipment	39, 96, 98	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; General environmental protection; Soil management and conservation; Waste management and recycling
	Waste disposal equipment	39	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; General environmental protection; Soil management and conservation; Waste management and recycling
	Incineration equipment	84, 85	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Waste management and recycling
	Recycling equipment	84	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Environmental goods and services promotion; General environmental protection; Waste management and recycling
Wastewater management	Water handling goods and equipment	73, 84, 90	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Plant protection; Soil management and conservation; Water management and conservation

Table 6: *Environmental objectives and OECD's environmental goods (continued)*

OECD category	OECD product type	Associated HS chapters	Harmonised environmental objective
	Aeration systems	84	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Soil management and conservation; Water management and conservation
	Oil/water separation systems	84	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Soil management and conservation; Waste management and recycling; Water management and conservation
	Screens/strainers	39, 84	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Soil management and conservation; Waste management and recycling; Water management and conservation
	Sewage treatment	58, 73, 84, 85	Biodiversity and ecosystem; Chemical, toxic and hazardous substances management; Environmental protection from pests and diseases; General environmental protection; Soil management and conservation; Waste management and recycling; Water management and conservation
Water supply	Potable water supply and distribution	22, 28, 39	Chemical, toxic and hazardous substances management; Climate change mitigation and adaptation; Water management and conservation
	Water purification systems	28	Chemical, toxic and hazardous substances management; Climate change mitigation and adaptation; Environmental protection from pests and diseases; Soil management and conservation; Water management and conservation

The key idea here is to assign a higher strength to the links for which the HS chapter corresponds to the activity described in the "harmonised sectors" and the "harmonised environmental objectives". This is idea is implemented by assigning a different weight to the links which are consistent with the economic sector and/or environmental objective associated with the measure.

To put it formally, let S_i denote the set of HS categories that match the "harmonised sectors" of measure i, and E_i be the set of HS chapters that are consistent with the "harmonised environmental objective" of measure i. Then we can introduce a weight W_{ij}^S and W_{ij}^E to adjust the link strength:

$$\tilde{L}_{ij} = L_{ij} \cdot W_{ij}^S \cdot W_{ij}^E \quad \text{with} \quad W_{ij}^S = \begin{cases} 1 & \text{if } j \in S_i \\ 0.5 & \text{otherwise} \end{cases}$$
$$W_{ij}^E = \begin{cases} 1 & \text{if } j \in E_i \\ 0.9 & \text{otherwise} \end{cases}$$

The weights are assigned based on our best judgment. Measures receive a weight of 0.5 if they do not match the "harmonized sectors", and a weight of 0.9 if they do not match the "harmonized environmental objective". The latter category of unmatched measures is given a higher weight to avoid creating an overly low link for policy measures that do not match either category. The weights also reflect our assessment that HS product codes correspond more strongly with broad economic sectors than with broadly defined environmental objectives.

Step 4: HS/ICS codes reported by members

Among the variables of the EDB, the "HS - ICS code" field is of particular interest. In 22% of the EDB notifications — primarily under the TBT agreement — members supplied the HS/ICS codes of the goods affected by the measure. This information can significantly simplify the matching of HS codes. In fact, for measures that come with product code information, we can restrict the search to the codes provided by the member. However, in order to use the product code information, there are two issues that we need to tackle:

1. Some of the product codes might refer to non-environmental measures notified by the member, therefore we need to identify the codes that are relevant to the environmental measure from the ones that are not;
2. ICS and HS codes are mixed in the notifications, therefore we need to find a way of recognising and converting ICS codes.

The first issue is tackled by considering the notified product codes as the *possible set* of codes for the measure. That is to say, any HS code matched to the measure must be among the ones reported by the member. Within this possible set of codes, the ones with the strongest links to the measure description are to be considered the most relevant to the environmental goal.

The second point requires more elaboration. ICS and HS codes are very similar, they are both numeric sequences of varying length, whose grouping is often (but not always) separated by dots. Their main distinctive features are the positioning of dots and the length of the second-level grouping, which is of 3 digits in ICS and 2 digits for HS. As a result, ICS tends to have an odd number of digits, while HS has an even number of digits. Building on this insight, we use a set of regular expressions to tell ICS codes apart from HS codes. An additional level of complexity is added by the fact that data may transit trough an excel spreadsheet. Whenever a notification reports only a single ICS/HS code, excel identifies the value in the cell as a number and will automatically remove leading and trailing zeros. The boxes below report the regexes used for measures that report multiple codes (top) and single codes (bottom) for HS and ICS codes.

```
HS:

    ^(\d\d\.?){2,6}$|^\d\.(\d\d\.?){1,5}$|^\d{3,4}\..*$|^\d{3,4}$

    ^\d?\d\.\d{4}.*$|^\d{3,4}\..*$|^\d{3,4}|^\d?\d\.\d{2}\..*$
```

```
ICS:

    ^(\d?\d)\.\d{3}(\..*)?$|^\d{5}(\d\d){0,2}$|(\d?\d)\.\d{5}$

    ^(\d?\d)\.\d{3}\..*?$
```

Essentially, these regular expressions identify the codes that are *exclusively* consistent with the pattern of ICS codes or HS codes. The next step, is to convert ICS into HS codes. There is no clear-cut conversion table. We rely on an internal conversion table developed by the WTO Secretariat along the same line of Han et al. (2019). The HS chapters obtained after the conversion form the *possible set* for the measure on which the link search is performed.

All the codes that are not unequivocally identified as HS or ICS are considered ambiguous. For example, any 2-digits code is ambiguous because it could either be an HS or ICS code. Another example would be any code of the type `15.8`; technically this is neither an HS nor an ICS code. The ambiguous codes are not discarded, they can still provide useful information. To every ambiguous code we match the closest possible HS and ICS code. For the example above, this would be the HS code `1580` and the ICS `15.800`. Then, the ICS code is converted to HS using the same conversion table. Finally, both the converted codes and the closest HS match are retained to define the possible set for the measure.

Step 5: Relative link strengths

As a next step, we express the link strength in relative terms, so as to have a measure that is comprised between 0 and 1 and reflect the probability of matching between measures and HS categories. For each measure, we calculate the relative strength $\bar{L}_{ij}$ of each one of its links:

$$\bar{L}_{ij} = \frac{\tilde{L}_{ij}}{\sum_{j=1}^{J^*} \tilde{L}_{ij}}$$

$\bar{L}_{ij}$ expresses for each measure i the relative strength of the HS category j according to our keywords matching.

Step 6: Reducing the number of links

The method presented so far gives rise to a high number of links. In fact, we find a total of 448637 links between measures and HS 2-digits categories. On average, this is 40 links per measure. A look at the distribution of $\tilde{L}$ reveals that the majority of the existing links have a low strength (see Figure 10). This suggests that many of the links are based on the matching of few generic words. Hence, we introduce three new parameters to tackle this problem:

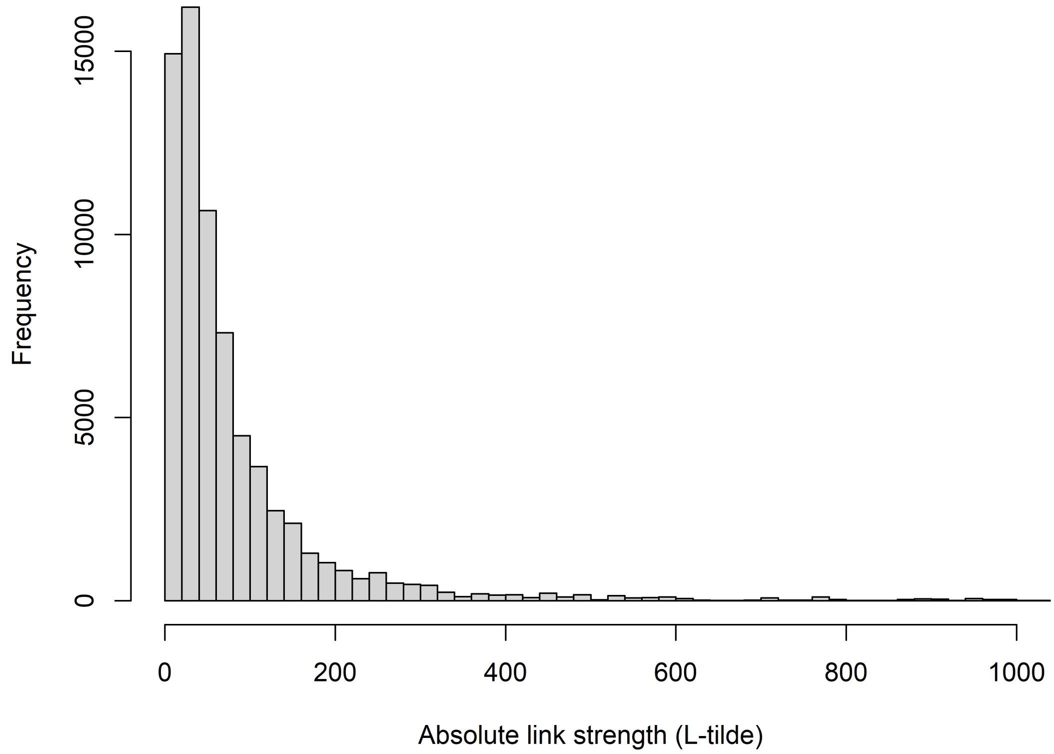


Figure 10: *Distribution of absolute link strengths ($\tilde{L}$)*

1. A first way of dealing with this problem is to filter the keywords used for matching. Since the high number of links derives from the matching of less-informative keywords, one could introduce a parameter that controls the minimum required keyword information. We implement this idea by setting a threshold value J^+ defined as the maximum number of HS categories in which keywords are allowed to appear. Then, the keyword weight ω_k of step 2 becomes:
$$\omega_k = \begin{cases} 1 + \log\left(\frac{1+J^+}{1+J_k}\right) & \text{, if } J_k \leq J^+ \\ 0 & \text{, if } J_k > J^+ \end{cases}$$
For example, $J^+ = 10$ would imply that all keywords that appear in more than 10 HS chapters are not used in the matching process. As a result, only the most informative keywords are used and the overall number of links is reduced.

2. Just like in Han et al. (2019), we also introduce a cut-off value for the absolute link strength to eliminate the weakest links. Let $\tilde{L}^+$ be the cut-off value for the absolute link strength.

Only the values above the cut-off are retained. This cut-off value is applied between step 4 and step 5.

3. In a addition to the above cut-off value, we also introduce a cut-off on the relative strength of links to be applied after step 5. Let $\bar{L}^+$ be the cut-off value for the relative link strength. Only the values above the cut-off are retained. This cut-off is effective at limiting the maximum number of links by measure. It particularly affects the measures that have been linked to a high number of HS chapters and, thus, have a more ambiguous match.

Step 7: Calibrating parameters and evaluating results

The value of the new three parameters are set in such a way as to minimise the average links per measure while maximising the number of measures linked. In order to get an understanding of the best values for the three parameters, we simulated the matching for different combinations of the three parameters. We then evaluated the matching performance by sampling a few measures and comparing the description of the measure and the matched HS score. We also compare the results of the matching with the HS/ICS codes provided under the TBT agreement and use this information to calibrate the cut-off points and keyword threshold of step 6. As we will now explain, the following values are selected:

$$J^+ = 20 \quad , \quad \tilde{L}^+ \approx 9.4 \text{ (70\% quantile)} \quad \text{and} \quad \bar{L}^+ = 0.1$$

The first parameters that is applied during the matching is the keyword threshold. By reducing the threshold, fewer and fewer measures are matched to HS codes because only the most informative keywords are kept. The keyword threshold value J^+ is only meaningful if set at stringent values (Figure 11). The threshold starts to become effective at reducing the total number of links only for $J^+ \leq 30$. It should be noted, that the effectiveness of this threshold increases almost exponentially as the threshold is reduced. However, the downside of setting an excessively low keyword threshold is that it might exclude keywords that are useful for matching, thus many EDB measures could be left unmatched. After analysing different threshold values and how they combine with the other parameters, we opted to set J^+ at the value of 20. As depicted in the top-left panel of Figure 11, this value is as low as it can get without causing significant reduction in the number of measures that can be matched. A value of 20 allows to keep sufficient keywords to potentially match up to 12850 EDB measures and should at the same time improve the quality of the matching by filtering out less informative keywords and ultimately reducing the likelihood of mismatches.

The second parameter applied to the data is $\tilde{L}^+$. For ease of interpretation, the value of $\tilde{L}^+$ will be reported as quantile of the distribution of $\tilde{L}$[10]. There is an obvious trade-off between the cut-off for the absolute link strength and the number of measures which are matched. Setting a cut-off value at the 70% quantile implies keeping only the 30% of the links with the highest strength. As shown in Figure 12, a first step is visible for low values of $\tilde{L}^+$. This step corresponds to the absolute weakest links. They are based on single words in the description of the HS chapter. Therefore it is important to set $\tilde{L}^+$ at least above this level. We decided to set the cut-off value at a high level (70% quantile) in order to take full advantage of the reduction in average links per

[10] For example, a value of $\tilde{L}^+ = 0.7$ corresponds to $\tilde{L} \approx 9.4$, for which 70% of the links have a value that is below the cut-off. It should be noted that the quantile of the distributions are affected by the keyword threshold. All the values reported in the text correspond to the quantiles obtained with the threshold value of $J^+ = 20$.

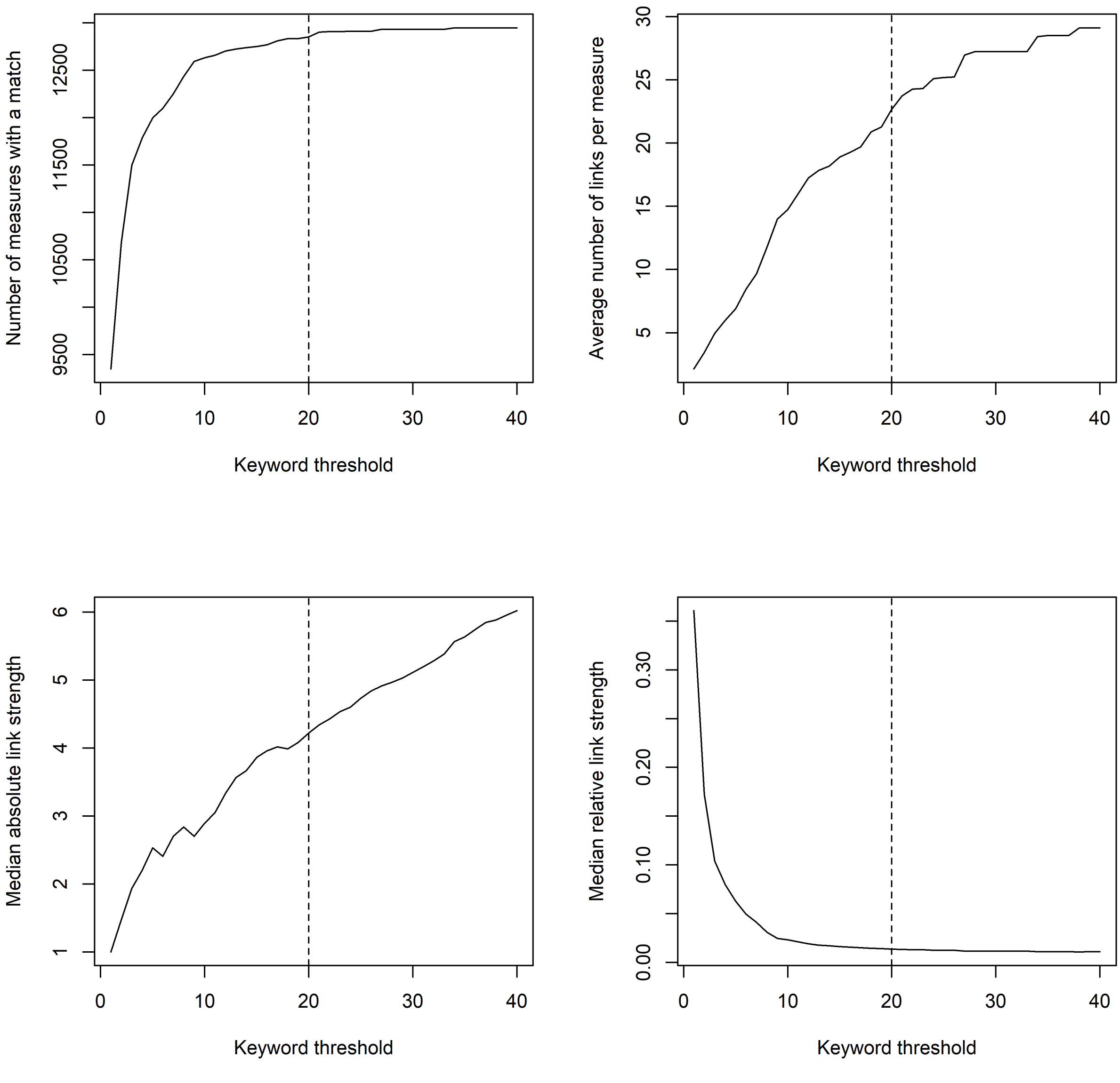


Figure 11: *Matching statistics as a function of the keyword threshold* J^+

measure while keeping the total number of matched measures relatively stable (top left and right panel of Figure 12).

Finally, the cut-off value on the relative strength of links is applied after the last step of the matching. Figure 13 depicts the number of measures matched (top left panel) and the average number of links per measure (top right panel) for increasing levels of $\tilde{L}^{+}$. Notice that for small levels of the cut-off there is almost no decrease in measures matched, whereas the average number of links per measure is significantly reduced. The reason is that the relative cut-off targets exclusively the links that have a lower matching probability for each measure. We take advantage of this by setting $\tilde{L}^{+} = 0.1$, i.e. only the links having a relative strength above 10% are retained.

After applying these three parameters, we are left with a total of 11123 measures linked to HS codes and an average of 2.7 links per measure. Figure 14 shows how frequently each HS chapter has been linked to environmental measures. As illustrated by the figure, chapter 84 and 85 attract a preponderant number of matches. Out of the 30487 links, 6507 are either to chapter 84 or 85. Besides these two chapters, we remark that chemical products are also frequently addressed by EDB measures.

We can better understand this result if we investigate the keywords used in the matching process. Table 7 shows the most frequent keywords used for matching in chapter 84 and 85. From these tables it appears that these two chapters match with some of the most common measure keywords. In particular "energy", "water" and "plant". Furthermore, it should be noted that chapter 84 and 85 are the two most common HS chapters in the OECD list of environmental goods. Chapter 84 covers "Nuclear reactors, boilers, machinery and mechanical appliances and parts thereof", while chapter 85 includes "Electrical machinery and parts thereof". They group a large and heterogeneous set of goods, many of which could be linked to sustainable agriculture and energy policies. For example, these chapters cover parts relating to engines (electric, combustion, etc...), turbines, purifying machines, photovoltaic panels, batteries and agricultural machinery.

To check the consistency of the results we tried: 1) to set J^{+} to 1, thus only keeping keywords that appear in a single HS chapter; 2) using only nouns and proper nouns for matching, that is to say excluding adjectives and verbs, both of which could be misleading out of context; 3) blocking some of the most frequent keywords of chapter 84 and 85 that do not appear directly linked to the goods covered by these chapters. Despite these attempts, the results remain stable: these two chapters consistently surpass all the others.

The Tables 8 and 9 below show respectively the top 5 and bottom 5 links by absolute link strength ($\tilde{L}$). Globally, the quality of the matches relies heavily on the length and character of the description of the measures. These descriptions do not follow a standardised template and they often do not detail the products affected. The wording is often generic and tends to relate to sectors of implementation rather than products. As a result, the matching with the HS classification may be unreliable at times. Nevertheless, in most cases, the matching is reasonably accurate at the 2-digits level. As shown in Table 8, the best matching is achieved when the coverage description includes a long list of products affected. However, such a comprehensive description is available only for a handful of measures. Conversely, the matching does not seem to perform well when the description is short and generic terms are used (see Table 9). Moreover, as discussed above, chapter 84, 85, and chemicals (28-39) attract a very high proportion of matches. In general, these chapters appear more often among the stronger links than the weaker ones.

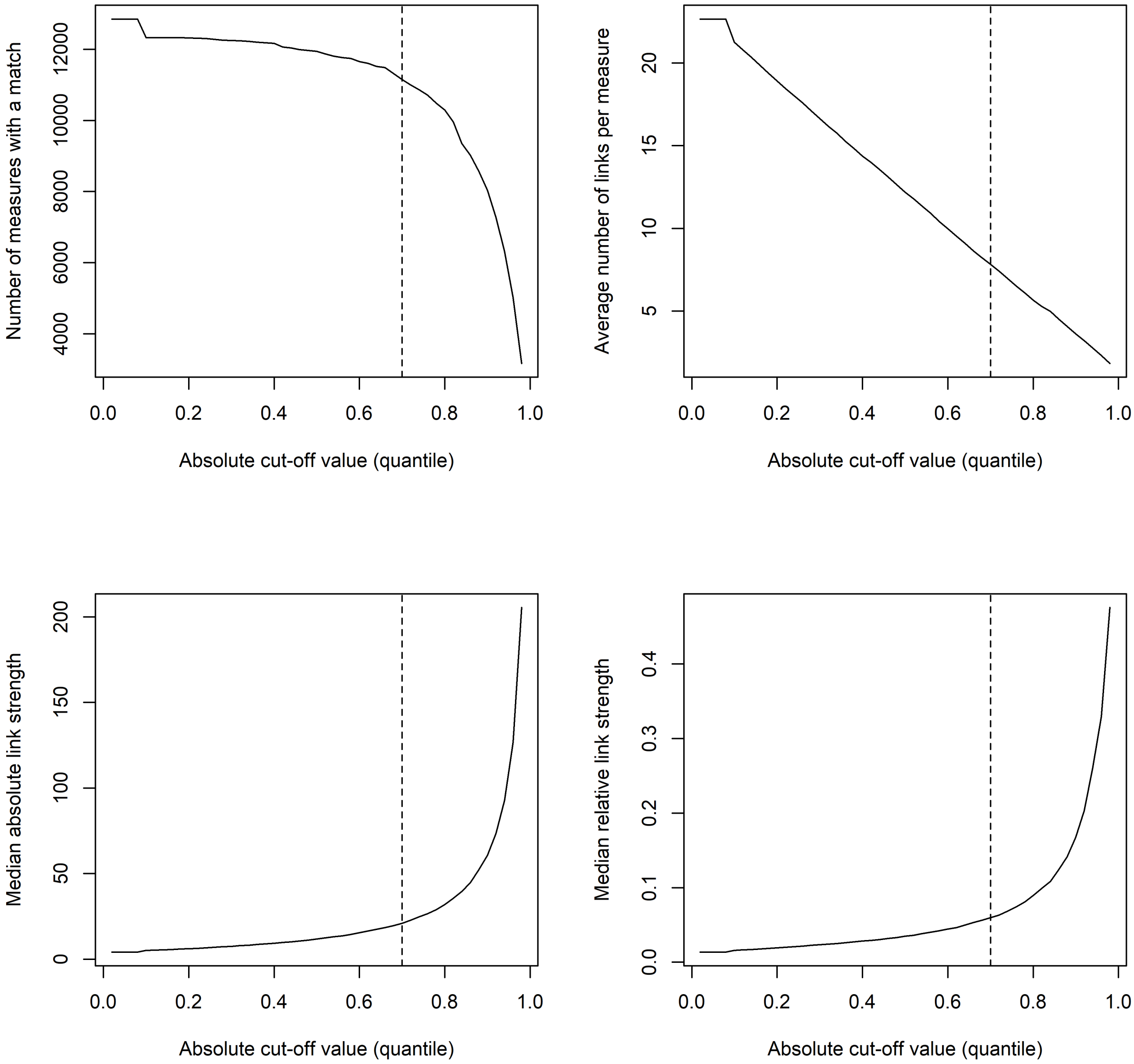


Figure 12: *Matching statistics as a function of the absolute cut-off value* $\tilde{L}^+$ *(with* $J^+ = 20$*)*

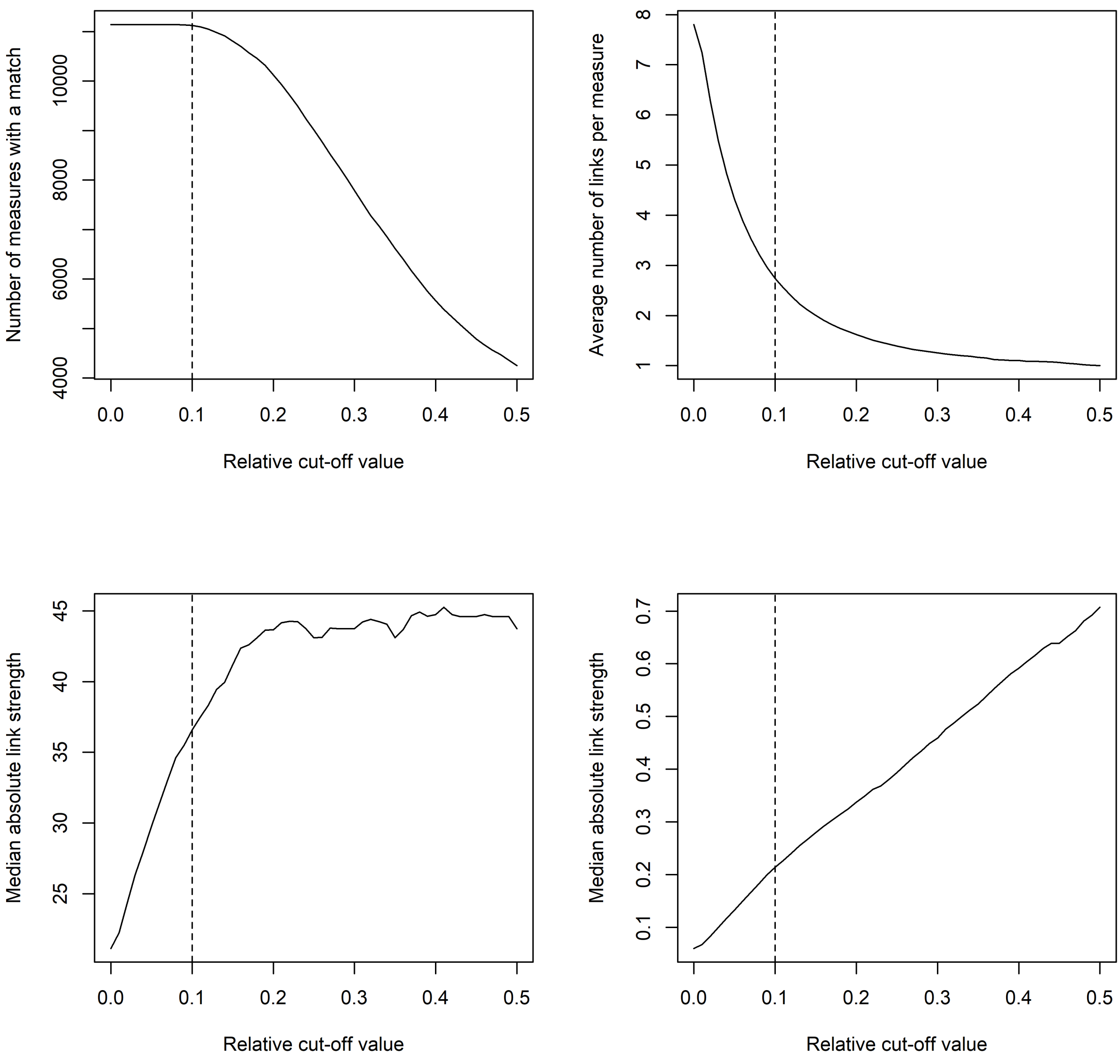


Figure 13: *Matching statistics as a function of different relative cut-off values* $\bar{L}^+$ *(with* $J^+ = 20$ *and* $\tilde{L}^+ = 0.7$ *)*

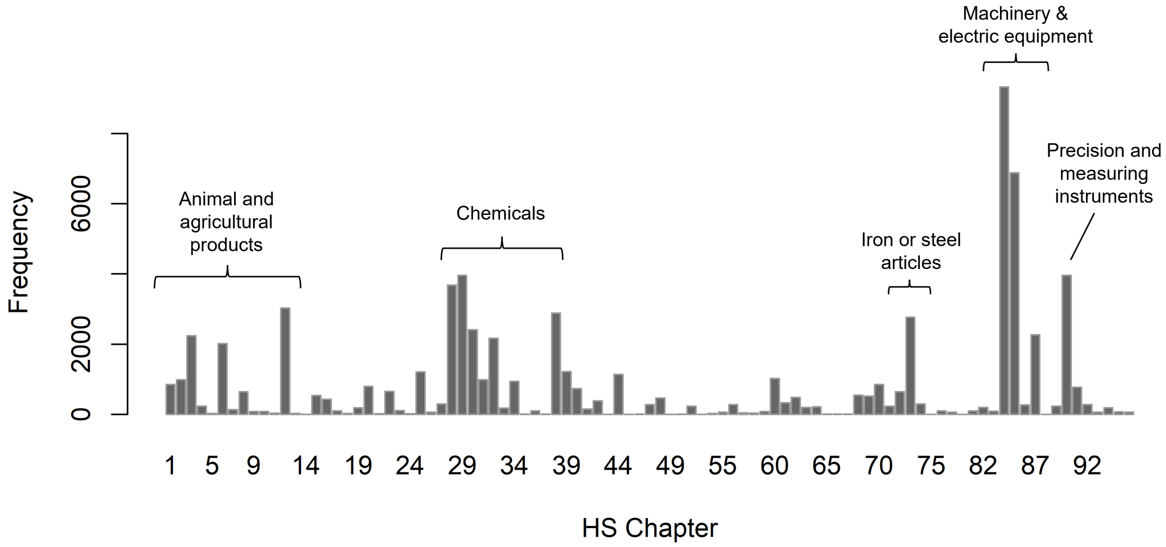


Figure 14: *Matching frequency of HS chapters*

Table 7: *Top 10 matching keywords*

All chapters		**Chapter 84**		**Chapter 85**	
keywords	**freq.**	**keywords**	**freq.**	**keywords**	**freq.**
natural	25940	general	1587	energy	2485
water	24264	plant	1440	water	1348
production	12042	water	1348	production	1338
plant	11520	production	1338	safety	1131
measure	11418	agricultural	1131	equipment	730
safety	10179	safety	1131	gas	651
gas	9765	measure	1038	food	608
human	9558	equipment	730	soil	602
general	9522	industry	700	air	535
equipment	8760	gas	651	industrial	521

Table 8: Top matches

Measure nr	Coverage description	HS chapters	HS description	$\tilde{L}$
10497	Granulated slag (slag sand) from the manufacture of ferrous metals; Slag, dross (other than granulated slag), scaling and other waste from the manufacture of ferrous metals; [...]	29	Organic chemicals	3819
3232	Compression ignition engines for vehicles, gas engines for vehicles, automobile vehicles spark-ignition reciprocating or rotary internal combustion piston engines. [...]	87	Vehicles, except railway or tramway, and parts	3824
76	Welding machine; Machinery and apparatus for soldering, brazing or welding, whether or not capable of cutting, other than those of heading 85.15; gas-operated surface tempering machines and appliances (HS 8468); Energy and heat transfer engineering in general	84	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof	4138
10491	Wastes, which composition includes as a component or contaminant any of the following substances: arsenic, arsenic compounds, mercury, mercury compounds (excluding mercury vapour lamps and fluorescent tubes); Magnesium dust; [...]	29	Organic chemicals	4286

Table 8: *Top matches (continued)*

Measure nr	Coverage description	HS chapters	HS description	$\tilde{L}$
11860	Hydrogen cyanide, Phosgene: Carbonyl dichloride, Phosphorus oxychloride, Phosphorus trichloride, Phosphorus pentachloride, Sulphur monochloride, Sulphur dichloride, Thionyl chloride, Cyanogen chloride, [...]	29	Organic chemicals	6962

Table 9: *Worst matches*

Measure nr	Coverage description	HS chapters	HS description	$\tilde{L}$
2927	Heat supply organizations	57	Carpets and other textile floor coverings	9
12274	[*no coverage description provided, matching based on the measure description column*]	22	Beverages, spirits and vinegar	9
12972	Manufacturing/processing and research/development projects	23	Residues and waste from the food industries; prepared animal fodder	9
6642	Eligible industries include clean energy technology	33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations	9
999	Government-invested research institutions, universities, research institutions and private enterprises that participate in the Environmental Technology Development Project	82	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal: Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal	9

C Similarity index

To help identify measures that have been renewed or notified multiple times, we calculate a similarity index between pairs of measures. The index is calculated based on the proportion of words in the measure descriptions that pairs of measures share in common. We start by tockenising the words in the description of the variables "Measure description", "Coverage of measure" and "Environment-related objective". Then we use the set of words extracted from each measure description to calculate the Jaccard index for any given pair of measures. For every pair of measures ij our similarity index S is given by the share of words that the two measure have in common, over the total number of unique words in the two sets:

$$S_{ij} = \frac{|W_i \cap W_j|}{|W_i \cup W_j|}$$

Where W_i and W_j are respectively the set of words of measure i and j. Given that the EDB contains more than 13000 measures, the number of ij combinations is extremely high (over 150 million). The calculation can be simplified by looking exclusively at pairs of measures which share at least one notifying member in common.

D Scoring policy measures

This appendix introduces an index of strength for the Environmental Database (EDB). The intention of this index is to proxy the regulatory strength of the enacted environmental measures, as captured by the notifications of the Members.

Policy measures are notoriously hard to quantify due to the many forms they can take and the difficulty in interpreting their economic implications. Subtle changes can have profound stringency implications, and the impact of a measure is highly specific to the country and sector in which the measure is implemented. Therefore, our index constructed from the EDB information can only capture part of the equation and should be used only as an indication of measure strength. This index does not constitute an official ranking of policies.

Given the multifaceted nature of environmental policies, we attempt to quantify the strength of EDB measures along two dimensions: the *breadth* and *depth* of the enacted policies (Figure 15).

- **Breadth**: The breadth of a measure is defined by the range of economic sectors and environmental issues that are affected by the policy. For example, a measure that limits the import of a specific pesticide used in corn plantations could be considered as a narrow policy measure. On the opposite, an economy-wide environmental tax could be considered as a broad policy measure because it affects a large proportion of the economy and might deal with multiple environmental issues. In the indicator proposed in this paper, breadth is measured by: 1) the share of the economy that is affected by the measure, 2) the number of environmental objectives pursued by the measure, and 3) the number keywords used for classifying the measure.
- **Depth**: The depth component refers to the intensity of the measure. This aspect is arguably harder to quantify with the EDB data. The proposed indicator of policy depth relies on: 1) the wording used in the description of the measure, environmental goal and measure coverage. 2) The variety of policy tools used under the measure — a measure with multiple tools is deemed stronger than a measure that relies on a single type of intervention. 3) The type of policy tool used in the measure. For instance, a ban or a tax are in general stronger than a quarantine requirement or a risk assessment.

D.1 Breadth component

The breadth component captures the scope of the measure in environmental and economic terms. Three indices are proposed here; they are all measured on a scale from 0 to 1 and capture a different aspect of policy breadth.

Economic sectors
A first measure of the breadth is based on the range of economic sectors affected by the measure. Our starting point is EDB's classification of the "harmonised types of sectors subject to the measure". Each measure can affect one or more of the following harmonised sectors: agriculture, chemicals, energy, fisheries, forestry, manufacturing, mining, services, other, all sectors/economic

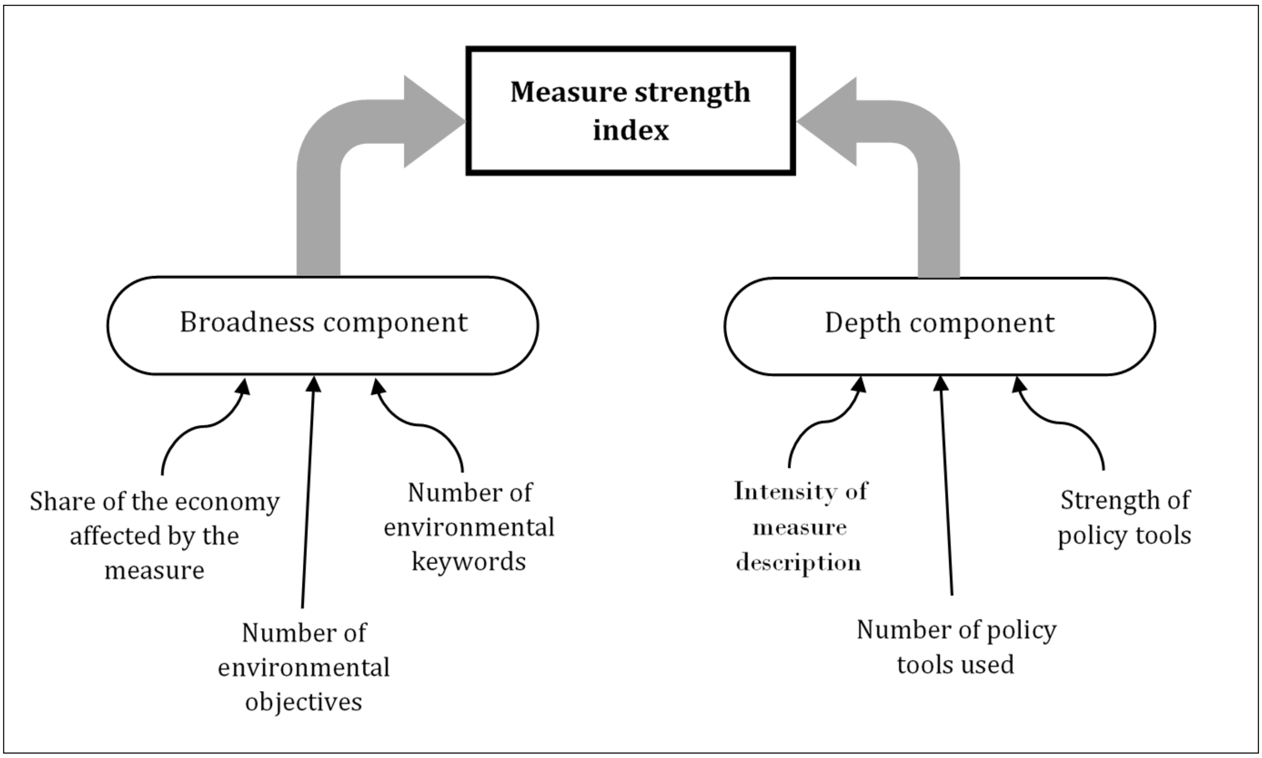


Figure 15: *Components of the measure strength index*

activities (and not specified).

The harmonised sectors give a good idea of the sectors affected by the measure, nonetheless, the importance of each sector might vary for different countries. So, for instance, if the economy of a country is predominantly based on the tertiary sector, an agricultural measure has a lower economic relevance than in a country whose economy is primarily based on agricultural production. We can take into account the subjective relevance of each sector by using national data on the share of value added by ISIC sectors.

The data on the economic share of each sector is taken from World Bank (2019) and UNSD (2020). To minimise the problem of missing data, we use the average over the period 2000-2018 as reference. In some cases, the data is not available for all sectors. We predict the missing data by regressing on the remaining available sectors accross the panel of countries (fractional logit). Finally, for a few sectors — such as forestry or fisheries — there is no available disaggregated data. We therefore assume they represent a constant proportion of the accounting unit in which they are included. For instance, the value added by fishing is assumed to be equal to one third of the value of "Agriculture, forestry and fishing" (ISIC group A), forestry is assumed to account for one sixth and agriculture for half of the value.

For every measure i of the EDB, an index of economic broadness is calculated as follows:

$$sectors_i = \frac{\log\left(1 + \sum_j^J h_{ij} \cdot S_{ij}\right)}{log(1 + 100)}$$

h_{ij} takes the value of 1 if the harmonised sector j is affected by measure i and 0 otherwise. S_{ij} indicates the share of harmonised sector j in the country of measure i. Essentially, we are calculating the share of the economy that is affected by the measure (which sums to 100). Since most of the measures affect a small share of the economy, we apply a logarithmic transformation to counterbalance the skew in the data and give more weight to differences in narrower measures. The denominator ensures the score is bounded between 0 and 1. A score of 1 indicates that all economic sectors are affected.

Number of environmental objectives

A second sub-component of breadth reflects the environmental ambition of the measure. The measure is considered broader if it tackles multiple environmental issues. We quantify this idea by counting the number of "harmonised environment related objectives" that are covered by the measure. Being a count variable, this sub-component follows a characteristic Poisson distribution. Therefore, we apply a logarithmic transformation to counterbalance the skew in the data and give a less-than-proportional weight to larger numbers of objectives.

$$objectives_i = \frac{\log(1 + E_i)}{\log(1 + max(E))}$$

E_i is the number of harmonised environment-related objectives of measure i and $max(E)$ is the maximum number observed in the EDB. Again, the denominator ensures the score is comprised between 0 and 1, where a score of 1 is assigned to the highest observed number of environmental objectives.

Number of environmental keywords

keywords is the last sub-component. This is another measure of environmental breadth based on the number of environmental keywords that have been used to tag EDB's entries. While environmental objectives describe primarily environmental goals (e.g. air pollution reduction, afforestation , etc.), environmental keywords describe areas of environmental policy (e.g. climate, energy, conservation, etc.). The calculation of this sub-component mirrors the method of the previous one:

$$keywords_i = \frac{\log(1 + K_i)}{\log(1 + max(K))}$$

K_i is the number of keywords of measure i and $max(K)$ is the highest number of keywords attached to a single entry of the EDB.

D.2 Depth component

The aim of a the depth component is to capture the intensity of the policy measure. Just like in the breadth case, the three sub-components are based on the variables of the EDB and each is measured on a scale ranging from 0 to 1.

Wording intensity

A first depth sub-component is based on the wording used in the description of the measure. Our goal is to assign a higher score to measures which have more assertive wording. To do so, we use the lemmatisation algorithm from `udipipe`[11] to extract all the verbs in their root form from

[11] The R package is available from `https://cran.r-project.org/web/packages/udpipe/index.html`.

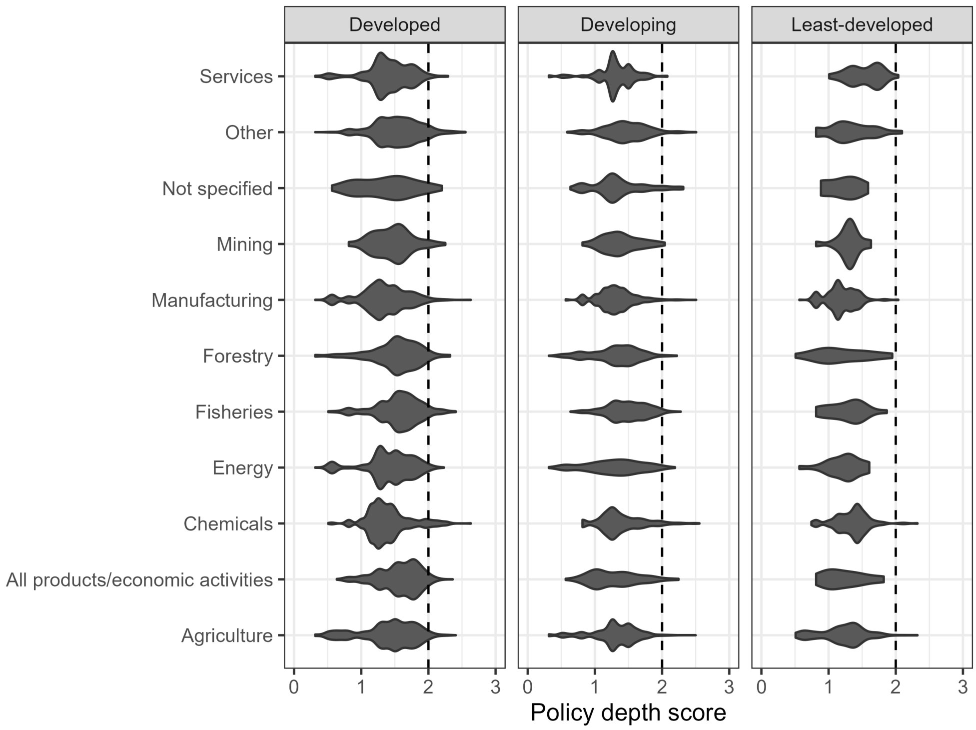


Figure 16: *Distribution of policy depth score by sector and countries' development status*

Notes: The figure illustrates the distribution of the policy score across sectors for countries with different development status. The vertical lines correspond to a score of 2, values above this threshold can be considered high. The figure shows that the policy depth score tends to be lower in LDCs. In particular, scores are the lowest in the energy and mining sectors.

the description of the measure, the description of the measure coverage and the description of the environmental objective of the measure. We then classify the 200 most frequent verbs according to their connotation in neutral, weak, average or strong. The table below shows the most frequent verbs in each group.

Table 10: *Verb grouping examples*

Neutral	**Weak**	**Average**	**Strong**
include	promote	protect	regulate
use	support	ensure	prevent
establish	contain	provide	require
propose	encourage	improve	prohibit
make	implement	reduce	exclude

We then devise a scoring system based on the frequency of verbs whereby the presence of stronger verbs is associated with higher scores. We first calculate:

$$W_i = \log(n_i^W) + 2\log(n_i^A) + 3\log(n_i^S)$$

where n^W, n^A and n^S indicate respectively the number of weak, average and strong verbs in the descriptions of measure i. The logarithms of the frequencies are used to give more weight to the first occurrences in each group of verbs. Then, the usual transformation is applied to bound the score between 0 and 1 and counterbalance the skewness.

$$wording_i = \frac{\log(1 + W_i)}{\log(1 + max(W))}$$

Variety of policy tools

A second sub-component of measure depth is based on the number of different policy tools that are adopted in the measure. We assume that the measure is likely to be stronger if multiple policy tools (e.g. grants, import quotas, regulation) are used. *variety* is calculated as follows:

$$variety_i = \frac{\log(1 + M_i)}{\log(1 + max(M))}$$

Where M_i is the number of harmonised types of measures identified for measure i. The usual logarithmic transformation is applied.

Measure types

The last depth sub-component is also built from the "harmonised types of measures" variable. Unlike *variety*, which looks at the number of different measures, *type* focuses on a tightness ranking of different policy tools. The ranking of policy tools is based on multiple characteristics, in particular, we regard as more stringent the measure types that are associated with higher compliance costs, are more direct and have a stronger coercive nature. Given the intrinsic variability within

Table 11: *Ranking of measure types*

Rank	Harmonised measure type
Standards and regulations	
1	Ban/Prohibition
1	Internal taxes
2	Import tariffs
2	Export tariffs
2	Import quotas
2	Export quotas
3	Technical regulation or specifications
3	Conformity assessment procedures
3	Import licences
3	Export licences
3	Services requirements
3	Quarantine requirements
3	Regulation affecting movement or transit
3	Environmental provisions in trade agreements
3	Other environmental requirements
4	Risk assessment
4	Countervailing measure / investigation
4	Intellectual property measures
4	Safeguard measure / investigation
4	Anti-dumping measure / investigation
4	Investment measures
Subsidies	
1	Grants and direct payments
1	Income or price support
2	Tax concessions
2	Loans and financing
2	Non-monetary support
2	Public procurement
2	Other price and market based measures
3	Other support measures
Other	
5	Not specified
5	Other measures

each measure type, we rank the measures in few broad groups. The ranking of each harmonised measure type is shown in the following table.

Each measure of the EDB is assigned the *type* score based on its highest-ranked measure type; measures in group 1, 2, 3, 4 and 5 are assigned respectively a score of 1, 0.75, 0.5, 0.25 and 0. For example, a measure that combines quarantine requirements with a ban/prohibition will be ranked in group 1 and given a score of 1. Then, the usual logarithmic transformation is applied:

$$variety_i = \frac{\log(1+T_i)}{\log(2)}$$

Notice that the denominator is log(2) because the maximum value assigned to measure type T_i is 1.

D.3 Details of calculation

For every measure i, the final strength score is obtained as a product of its depth component and breadth component.

$$\boldsymbol{Score}_i = Breadth_i \times Depth_i$$

Where *Breadth* and *Depth* are two components obtained by summing all the sub-components presented in section D.1 and D.2:

$$\begin{aligned} Breadth_i &= 1.5 \cdot sectors_i + 0.75 \cdot (objectives_i + keywords_i) \\ Depth_i &= wording_i + variety_i + type_i \end{aligned}$$

The final strength index, *Score*, is expressed on scale from 0 to 9 and is obtained by multiplying the *breadth* and *depth* components presented above. Both *Breadth* and *Depth* range between 0 and 3. Weights are applied to the indices in *Breadth* so that the contribution of *sectors* accounts for half of the breadth measure and the other half is determined by the environmental broadness captured by *objectives* and *keywords*.

Figure 17 shows the distribution of the policy score for *REG* and *SUB* variables. We see that *SUB* measures have on average a higher policy depth score, however a larger number of *REG* measures remain in force.

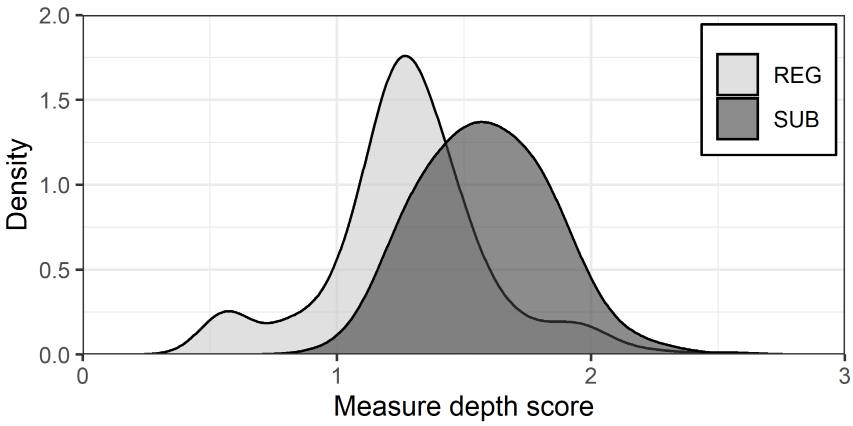


Figure 17: *Measure score distribution for two groups of policy measures*

The single and joint distribution of the two components are illustrated in Figure 19. The two components, as estimated in this note, are highly uncorrelated. This suggest that there is little overlap in the dimensions captured by these metrics.

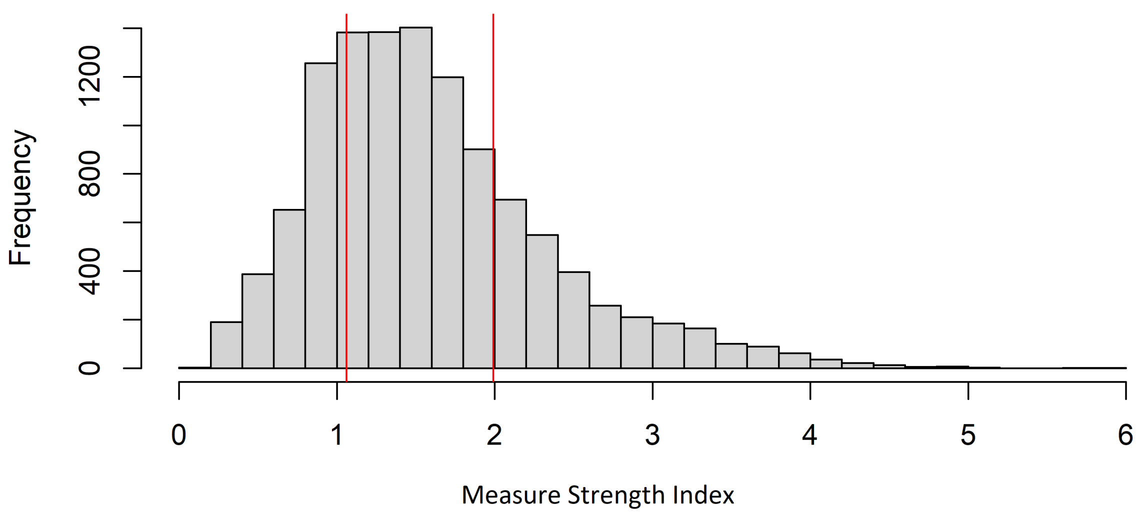


Figure 18: *Distribution of the composite index of measure strength*

The red lines indicate respectively the first and third quartile of the distribution. That is to say, approximately 50% of the EDB measures have a score between 1 and 2.

The score is measured on an abstract scale. Hence, it does not possess a direct numerical interpretation. As a rule of thumb, we could say that any measure with a score higher than 2 could be regarded as a "strong" environmental measure. In fact, approximately 50% of the measures have a score comprises between 1 and 2 (see Figure 18), which could be interpreted as an average score. Measures with lower score values are expected to have a weaker environment impact and be characterised by the use of less coercive policy tools. Among all the measures in the EDB, the lowest score is 0.18, and the highest is 5.81. Extreme values (above 6) are very hard to register since they would entail a measure that is extremely broad and stringent at the same time. As a reference, the following table lists the 3 measures with the highest and lowest score.

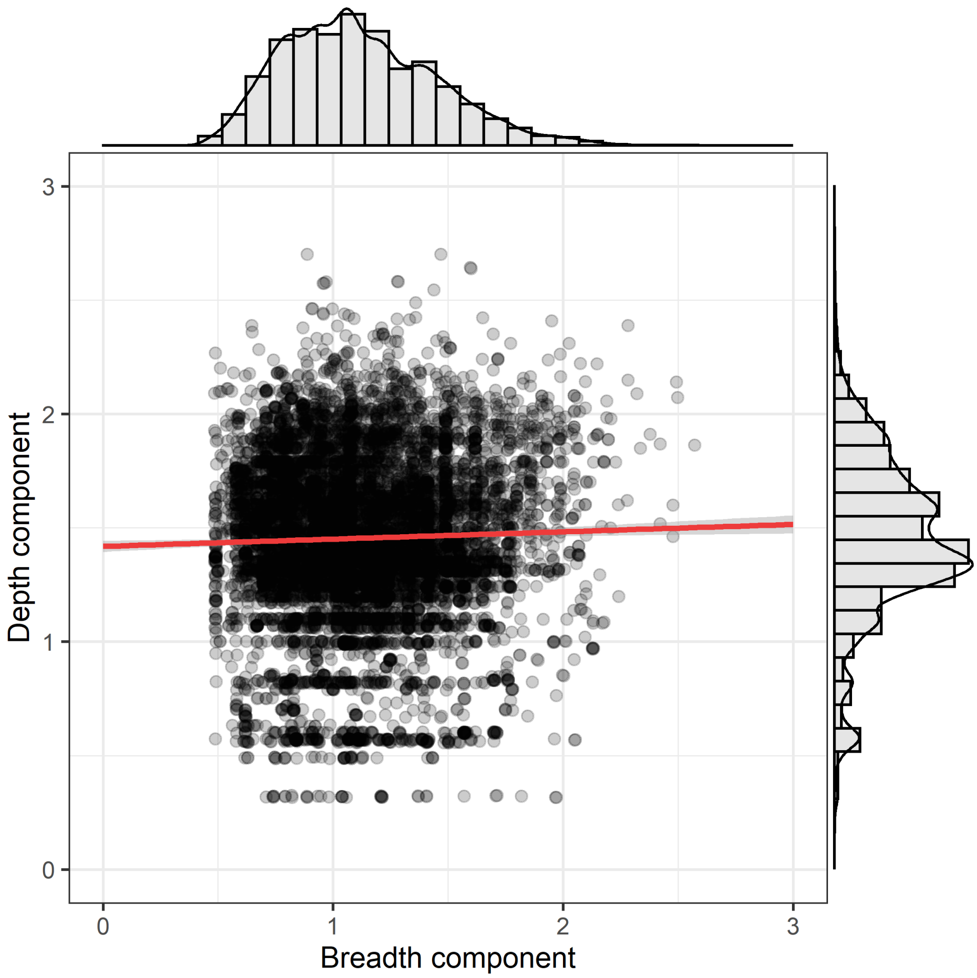


Figure 19: *Marginal and joint distribution of the depth and breadth components*

Notes: The distributions on the sides of the chart are respectively the breadth component (top) and depth component (right) marginal distributions. As illustrated by the flat red fitted line, the correlation between the two score component is extremely low.

Table 12: *Top and bottom 3 measures by strength index*

Nr	Agreement	Member	Keywords	Type of measure	Sectors	Strength index
3578	Agriculture	Canada	Environment; Conservation; Sustainable; Bio; Climate; Soil; Pollution; Natural resources; Wildlife	Grants and direct payments; Non-monetary support	Agriculture; Services	4.72
4589	SCM	Malta	Waste; Environment; Bio; Hazardous; Sustainable; Natural resources; Energy	Grants and direct payments; Loans and financing; Tax concessions	All products/economic activities	4.64
2062	SCM	Lithuania	Fish; Environment; Bio; Energy; Conservation; Climate; Renewable; Eco; Wildlife	Grants and direct payments	Energy; Fisheries; Services	4.53
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9668	Agriculture	Australia	Environment	Not specified	Agriculture	0.99
11294	Agriculture	Norway	Environment	Not specified	Agriculture	0.94
11295	Agriculture	Norway	Environment	Not specified	Agriculture	0.94

Figure 20 illustrates the score distribution for measures in different economic sectors. Considering both the breadth and depth of the measure, one can observe a high degree of variability within each sector (left panel of Figure 20). As shown by these distributions, our score definition emphasises the economic impact of the measure by giving higher scores to measures that affect a larger portion of the economy. Alternatively, the right panel of Figure 20 shows the score distribution of the depth component alone. The figure displays less variability across sectors.

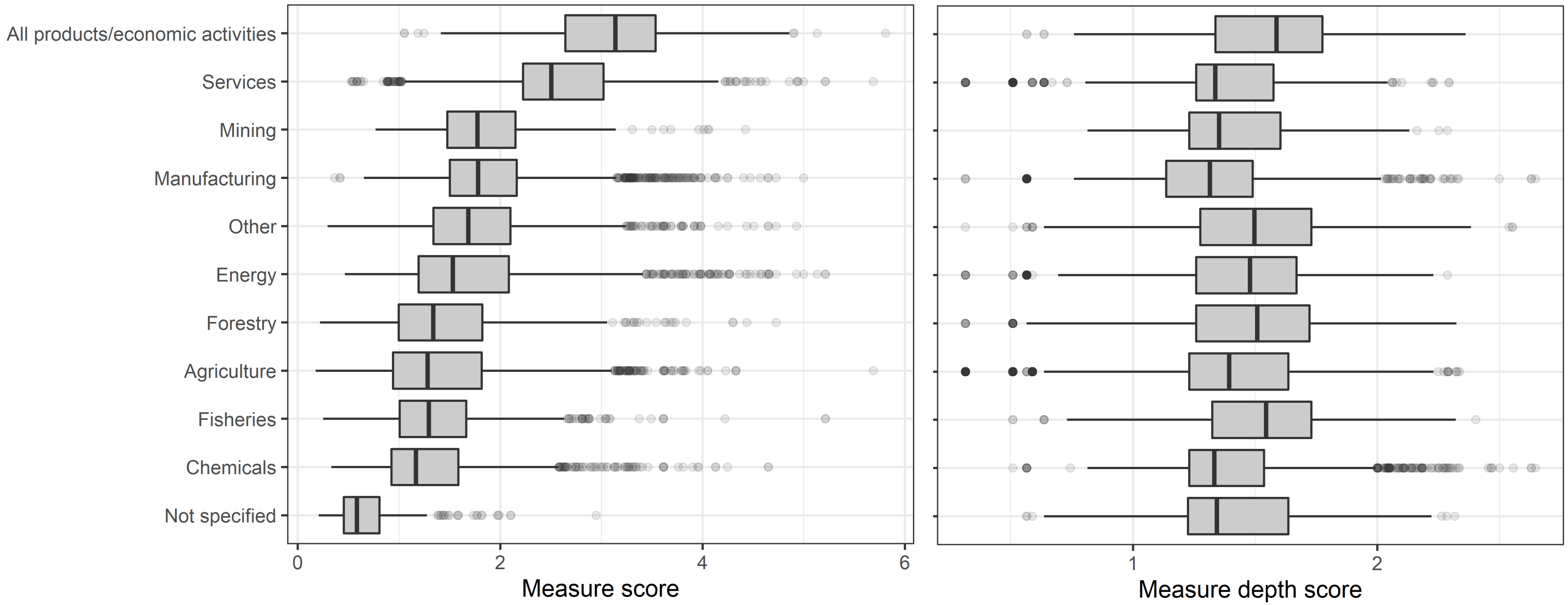


Figure 20: *Distribution of measure score by sector: total score and depth component*

Notes: The black vertical ticks are the median scores for each sector. The grey boxes encapsulate the first and third quartiles of the score distribution.

E Data sources and description

Patent data Data on the number of patent by IPC subclass code (e.g. `A01P`) comes from the OECD patent dataset (OECD, 2020). Only patents in the "triadic family" — a subset of patents filed both at the USPTO and EPO or JPO — are taken into account in order to exclude minor innovations from the sample. In fact, lesser innovations are usually not worth the higher cost of patenting in multiple jurisdictions. The "Triadic" definition is more stringent than patents with Patent Co-operation Treaty (PCT) application, therefore it selects higher-quality patents (OECD, 2009). We take the priority date (date of application in the first patent office) as date of reference for the innovation and consider it took place at the inventor's country of residence. The variable is fractional because the inventors could be based in multiple countries. The geographical coverage of the dataset is limited to around 110 countries, which is less than the trade and environmental measure data. The knowledge stock by IPC code is calculated by cumulating the number of patents from 1985 to year $t-1$ and depreciating it at a 15% yearly rate. To ease interpretation of the regression coefficients and result tabulation, the knowledge stock is expressed in tens of thousands of patents.

Trade data Trade flows at the 6-digits HS level (HS 2007 classification) come from the BACI dataset (Gaulier & Zignago, 2010). Earlier trade data is converted to the HS2007 nomenclature by using the concordance tables developed by the UNSD.[12] The BACI dataset is based on Comtrade data (UN, 2020) In the original dataset the trade flows of France and Monaco, Switzerland and Liechtenstein, and Belgium and Luxembourg are aggregated. We impute all the trade to the major of the two countries — thus treating Monaco, Liechtenstein and Luxembourg as *NA*.

Environmental measures All information on environmental measures comes from the Environmental Database (WTO, 2020). Refer to section 2 for more details. Each measure is linked to one or more HS 2-digits code based on the wording of measure descriptions (see Appendix B). The date of implementation of each measures is extracted via automated text analysis from the EDB (see Appendix A). Whenever it is impossible to determine the initial year of implementation, it is assumed that the implementation starts on the year of notification. Unlike subsidies, regulation measures are assumed to have no end date.

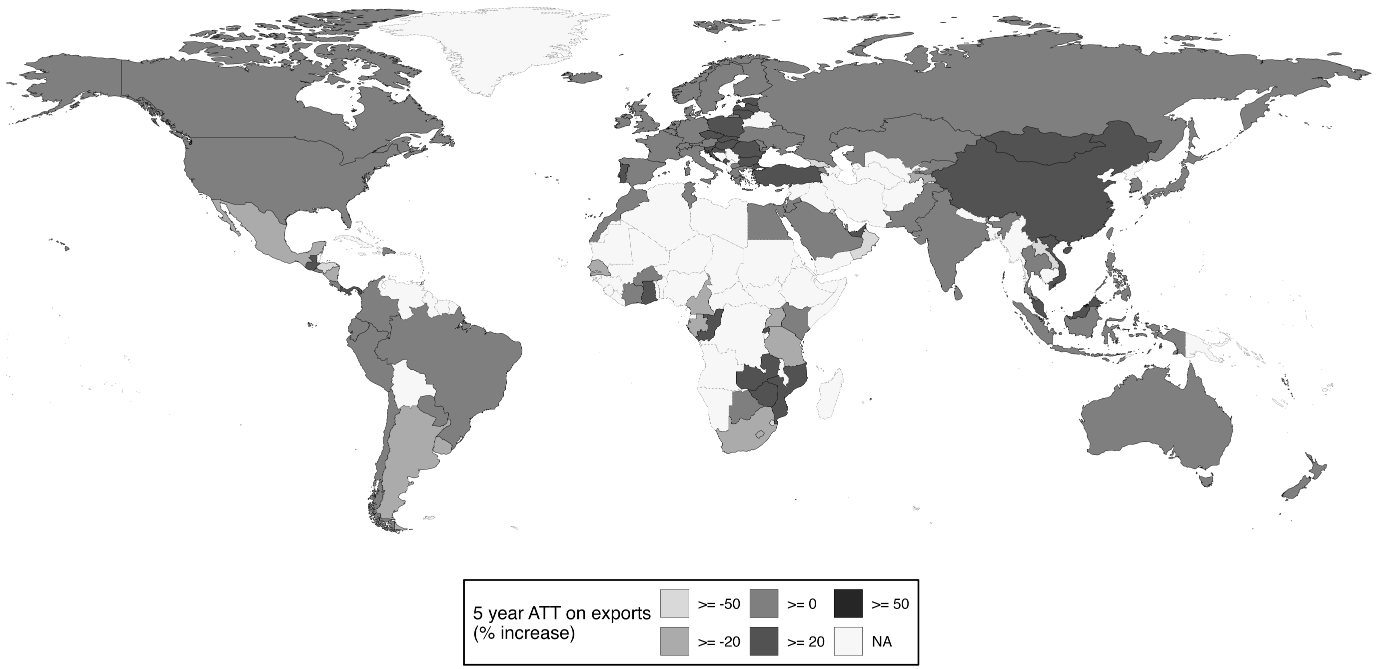
Environmental HS codes Environmental HS codes are identified with the OECD Combined List of Environmental Goods (Sauvage, 2014). The list contains 161 HS 6-digits codes that are related to the environment. These are all categories of goods that are related to environmental objectives such as air pollution control, water management, environmental monitoring or renewable energies.

HS – IPC – ISIC concordance Lybbert & Zolas (2014) developed a set of concordance tables between multiple versions of the HS, ISIC and IPC classifications. These tables are used to match the HS codes that are relevant to each IPC codes. The tables link IPC subclasses (e.g. B01D) of the 2006 revision to the HS 6-digits codes of the 2007 HS classification. The versions of the classifications are chosen to match the ones used in the trade and patent data.

[12] https://unstats.un.org/unsd/classifications/Econ

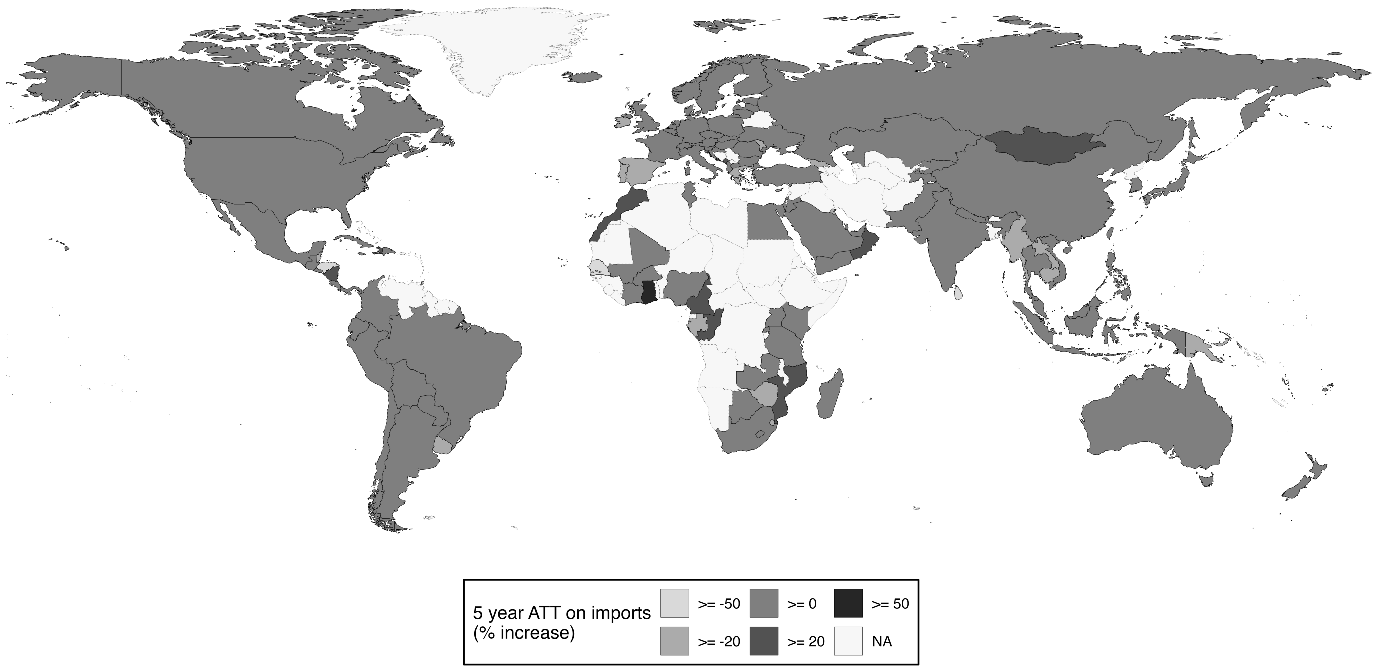
F Additional results

Figure 21: *Potential effects of environmental policy: aggregated export effects by country*



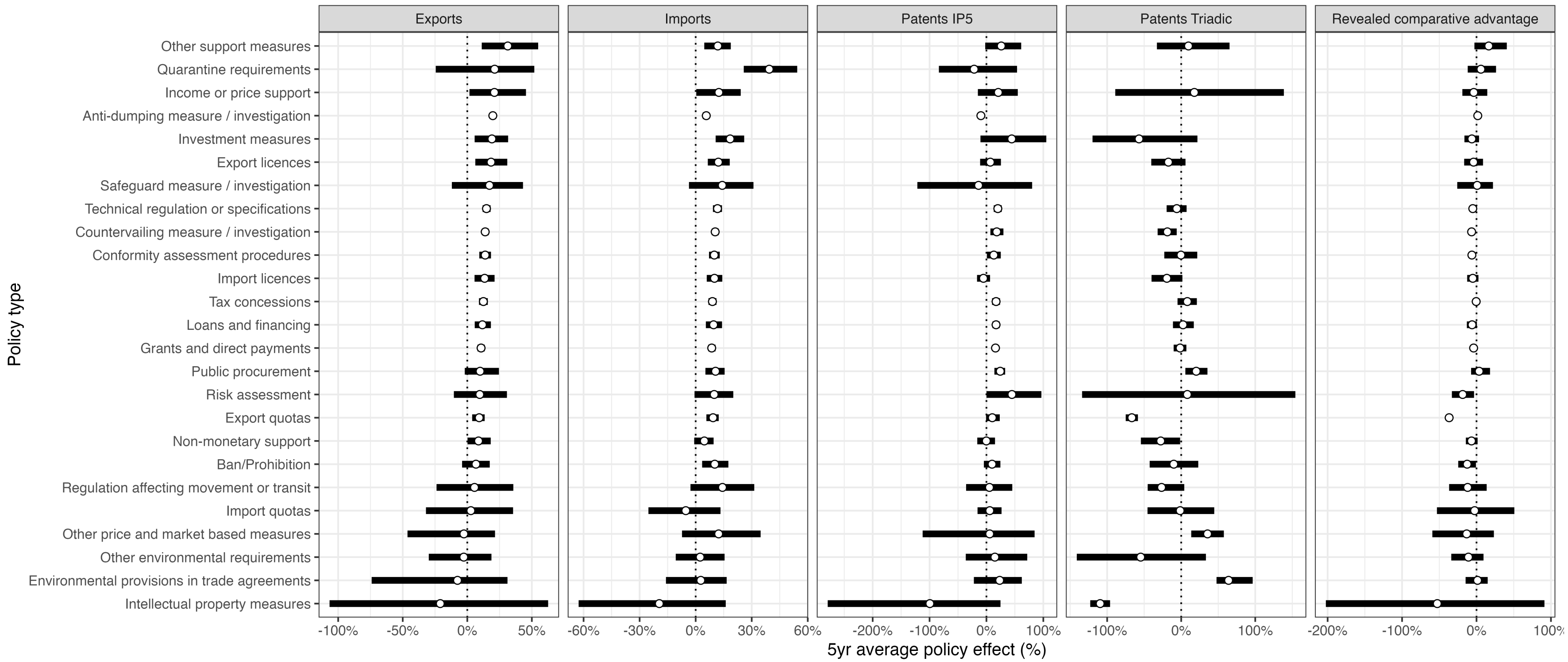
Notes: The plot displays the five-year average treatment effect on the exports of a treated country following the implementation of an environmental policy.

Figure 22: *Potential effects of environmental policy: aggregated import effects by country*



Notes: The plot displays the five-year average treatment effect on the imports of a treated country following the implementation of an environmental policy.

Figure 23: *Potential effects of environmental policy: individual measures*



Notes: The plot displays the five-year average treatment effect on a treated country relative to non-treated countries in a sector, following the implementation of an environmental policy.