# AUTOCATALYTIC NETWORKS AND THE GREEN ECONOMY

*Documents de travail GREDEG*
*GREDEG Working Papers Series*

Arnaud Persenda
Alexandre Ruiz

**GREDEG WP No. 2023-16**
https://ideas.repec.org/s/gre/wpaper.html

# Autocatalytic Networks and the Green Economy

Arnaud Persenda[1] and Alexandre Ruiz [*,1]

[1]Université Côte d'Azur, CNRS, GREDEG, France

GREDEG Working Paper No. 2023–16

**Abstract**

Green goods are products necessary to reach sustainable development targets. Since they offer many benefits, we discuss the question of why not all countries produce them. Using the economic complexity framework, we study how likely it is that a country will get a comparative advantage by producing green goods. We also study the externalities in terms of diversification prospects that arise from gaining a comparative advantage by producing and exporting green goods. For this purpose, we define a directed network in which nodes are products and links are the probability that a product catalyzes another one several years later. This network uses bilateral trade flows at the 6-digit level to assess the autocatalytic structure of product adoption, by identifying clusters of self-reinforcing products. We show that green goods are less prone to self-reinforcement compared to their non-green counterparts and offer fewer avenues for economic diversification. We also find that the impact of diversification varies across countries, suggesting that a one-size-fits-all approach to fostering the production of green goods may not be effective.

Keywords: Economic complexity, Economic growth, Structural change, Networks, Autocatalytic set

JEL classification: D85, F43, O25, O44, O50, Q50, Q56

*Corresponding author: `alexandre.ruiz@univ-cotedazur.fr`

# 1 Introduction

A vision of "green growth" has become the dominant policy response to climate change, supported by international organizations such as the OCDE (2011), United Nations (2012) and World Bank (2012). The main idea is to achieve global economic growth while at the same time reaching environmental targets. However, the scientific debate remains open as to whether economic growth is compatible with curbing climate change and keeping the global temperature rise below $2^o$C, as set out in the Paris agreement (UNFCCC, 2015). Some researchers argue that we can maintain global economic growth within planetary boundaries (see Capasso et al., 2019, for a summary), while others remain skeptical about the feasibility of this goal (Jackson, 2009; Hickel, 2018; Hickel and Kallis, 2020).

Policy intervention is a key factor in the debate on how to achieve climate goals. From a theoretical perspective, environmental and technology externalities should be addressed by adequate public policies (see Jaffe et al., 2005; Fischer and Newell, 2008; Popp et al., 2009; Acemoglu et al., 2012). In addition to allowing faster transition and preventing environmental disasters, Green Industrial Policies (GIPs) are gaining popularity as they support national competitiveness and job creation (Fankhauser et al., 2013; Rodrik, 2014). Furthermore, the notion of "green growth" sheds new light on aspects of structural change (see Savona and Ciarli, 2019, for a review) that affect greenhouse gas emissions.

GIPs are policies concerning the use of so-called green goods. Green goods are products or technologies that contribute to environmental sustainability through one or more of the following ways: reducing pollution, promoting resource efficiency, supporting environmental monitoring and protection, enhancing energy management, and facilitating sustainable end use or disposal.

Our paper contributes to this field by investigating the interdependences between green and non-green industries. Using a product space modeling, we model four distinct cases that reflect the interdependence between the two types of industries. The first case pertains to the potential for polluting sectors to emerge as a result of production capabilities acquired in a green sector. The second case involves a situation in which developing production capabilities in a non-green sector augments the probability of a green sector's emergence. The third case delves into instances where at least two sectors, one green and one non-green mutually reinforce each other, leading to the development of positive feedback loops that mutually benefit both sectors. Finally, the fourth case examines situations where feedback loops operate exclusively within green sectors or within non-green sectors.

These dynamic interactions shape the economic and environmental landscape, and understanding them is crucial for policy-making and industrial strategy. A key issue that has been overlooked in the literature is that green products are seldom completely green, as green industries are not isolated but are instead embedded in a value chain together with several other industries that have more or less environmental impact. Our paper addresses this point by investigating the interdependences between green and non-green industries.

Green goods are mostly supplied by industrialized countries. If developing countries want to ride the wave of green growth, they need to foster their own production capabilities. Such production bolsters environmental protection and resilience against climate change, a concern that is increasingly significant for many low and middle-income nations. This is particularly true since some low and middle-income nations have limited access to international resources and supply chains. Moreover, producing green goods often requires the adoption and mastery of advanced technologies. This encourages local innovation and technical skill development and catalyzes the transfer of technology from high-income to lower-income countries, thus promoting broader economic and social development. Finally, producing green goods domestically reduces import dependence, improving trade balance and economic resilience.

A burgeoning literature has developed on the subject of GIPs that aims to uncover stylized empirical facts about industrial capabilities (i.e., inputs, technologies and tacit knowledge needed to produce a specific good), competitiveness and environmental performance. Based on the economic complexity framework presented by Hidalgo et al. (2007), and Hidalgo and Hausmann (2009), researchers use a country's export portfolio to gain insights into its production capabilities. Some of those insights include (i) level of sophistication of the goods exported by the country, (ii) level of sophistication of the country's production system and (iii) diversification prospects at the product level.

Applied to green industrial development, this framework shows that GIP are more effective if the country already has capabilities (Huberty and Zachmann, 2011). Barbieri et al. (2023) show that new green technologies are often the recombination of existing technologies, including non-green ones. Recently, Fraccascia et al. (2018), Lapatinas et al. (2019), Mealy and Teytelboym (2020) and Romero and Gramkow (2021) have provided strong evidence that economic sophistication contributes to high capabilities in environmentally friendly production - or green goods - and reduces greenhouse gas emissions. The contribution of this paper is twofold. First, following methods introduced by Napolitano et al. (2018) we create a directed product space providing a dynamic and empirical overview of the positive externalities between each product produced by a given country. There is a positive externality between two goods when the production of one good significantly increases the probability of starting to produce the other good years later. Our results suggest that green goods are, on average, less likely to create production capabilities in other goods over time compared to non-green goods. Second, following Jain and Krishna (2003), we classify goods according to their importance in the product space by using a network approach based on the key concept of an autocatalytic set, which aims to capture and classify how products benefit and contribute to each other. More precisely we divide the product space into: the core, where there is a strong self-reinforcing cyclical catalytic structure; the periphery, which is a set of products catalyzed by the core but which do not catalyze it in return; and the rest of the network. On the basis of this classification, we show that the structure of the autocatalytic set (ACS) is relatively stable over time, allowing us to identify a pattern in the reinforcement between green goods and other non-green and green products. Based on those two contributions we derive some key industrial policies.

The rest of the paper is organized as follows: Section 2 presents our literature background. Section 3 describes the methodological framework and the data. In Section 4 we discuss the results. Section 5 concludes.

# 2 Literature background: Theory and empirical evidence

A greener economy relies on specific technologies, some of which are embedded into green goods. Green goods are defined as products essential for sustainable development. Understanding how to foster the emergence of the industries necessary to produce green goods is therefore important. In this paper, we aim to uncover the underlying mechanisms that limit the ability of some countries to acquire production capabilities in environmental products. We also investigate if these mechanisms can explain situations in which investing in green industries leads to the emergence of more polluting industries. To answer this question we rely on the theoretical framework from Hausmann's Capability Theory. This framework emphasizes that a country's economic growth and diversification process is driven by its existing capabilities (see Hausmann et al., 2013). To produce a given good a country requires the combination of a specific set of capabilities. Capabilities in this context are non-tradable inputs necessary for the production process. These capabilities include infrastructure (such as cargo airports, container ports, railroads, pipelines, and high-speed internet), resources (such as oil reserves and arable land), institutions (such as laws, property rights, and education), and collective knowledge (Ortiz-Ospina and Beltekian, 2018).

Countries have different sets of capabilities. Some countries have a large set of capabilities and are therefore able to create many products, including sophisticated ones that require a combination of many capabilities. Other countries have a small set of capabilities and so cannot produce the most complex goods. Following this approach, we make the assumption that a country only makes the products for which it has all the required capabilities. To start the production of environmental goods, a country must therefore acquire the necessary capabilities. Capability accumulation is a necessary step toward economic development and green industrial development. Decision-makers must, however, be aware that on the path to the acquisition of the capabilities for environmental goods, it may also acquire the capabilities for polluting industries. Indeed, acquiring those capabilities could reduce the cost for the establishment of polluting industries and increase the likeliness of their emergence. Reciprocally, the capabilities used to produce sophisticated green goods might be found in less sophisticated industries. A decision-maker might need to acquire production capabilities for those simpler productions first to develop the capabilities to produce more sophisticated products later.

This theory suggests that countries should leverage their existing strengths and capabilities to create environmentally friendly products. For example, if a country has a strong manufacturing sector with expertise in producing a large set of turbines, it could build upon that capability to develop a manufacturing industry in wind turbines. Similarly, if a country has a skilled workforce in a specific sector such as the textile industry, it could use that expertise to develop greener alternatives.

The aim of this paper is to model these interconnections between green and non-green goods. "Product relatedness," introduced by Hausmann and Klinger (2006), models the interrelation between products based on their probability of co-occurrence in a country's export basket. Formally, goods (a) and (b) are related if both products are likely to be exported together. The product space framework is explained in detail in Appendix A.2.1. Thus, exporting (a) is expected to increase the probability of exporting (b) and vice versa. The "product space" is the network modeling of this list of pairwise interrelations (Hausmann and Klinger, 2006; Hidalgo and Hausmann, 2009). The product space is defined as a network model in which nodes are products and links are their relatedness. These spaces have been used to model product relatedness in order to understand economic development. Hausmann and Klinger (2006) contribute to the more general industry relatedness literature (see Teece et al., 1994) by introducing the question of economic development and regional product diversification.

To have an accurate measure of what a country produces, we follow the literature by relying on export data as a proxy. These data have several advantages (i) they cover most countries, (ii) they provide a reliable and extensive list of goods produced and (iii) databases are available over a large time span. This broad coverage makes it possible to measure production capabilities for most countries and territories, including those that do not have reliable production statistics. More specifically, in the context of the product and industry spaces used in the literature, the use of trade data is a trade-off between data availability and data coverage. The use of export data consists of observing only a small subset of global economic activity which are the goods exported by a small group of competitive firms. However, the wide coverage, the standardization and the high level of disaggregation make it one of the best proxies to study industrial development, especially since it includes extensive information on low-income countries that may not have reliable production statistics.

The contribution of Hausmann and Klinger (2006) was to model those relatedness relationships in a network called the product space. In this network, nodes are products and links are the probability of two products to be co-exported[1].

In Hausmann's product space, there is no explicit explanation of how goods relate to each other. Moreover, links measure proximity between pairs of goods as a proxy for shared capabilities. However, this model does not offer a way to determine if the existence of a particular product increases the likelihood of another product's presence. Such relationship should exist. According to Hausmann's Capability Theory a good is produced when a country has all the required capabilities. Products differ in their required capabilities and therefore some goods require more capabilities than others. The more capabilities a product requires, the more sophisticated it is. Therefore some goods can be the recombination of capabilities used in other products. In consequence, if a product (a) requires the capabilities (A,C,D) and the product (b) requires the capabilities (A,B,C,D), then a country that produces (b) is likely already producing the product (a) as it has all the required capabilities. If this is not the case, the country can easily start producing (a) in the short term. Conversely, building the capabilities to produce good (a) could be useful in order to achieve production of (b). For policy-makers, having information on the hierarchies among products is crucial in order to devise efficient diversification policies.

Several attempts have been made to establish relationships between economic activities, both at the industry and product level[2]. The earliest directed product space, constructed using export data, was the "taxonomy matrix" proposed by Zaccaria et al. (2014). This network operates on the principle that the production of

[1]The authors rely on the uncomtrade database (Feenstra et al., 2005) in order to model the evolution of the product space from 1962-2000 at Standard International Trade Classification (Revision 4) at the 4-digit level. The proximity between goods (a) and (b) is the minimum of the number of co-exportations of goods (a) and (b) normalized by the ubiquity of good (a) (the number of countries that export good (a) (with a RCA $> 1$)) and the number of coexportations of goods (a) and (b) normalized by the ubiquity of good (b).

[2]For a previous instance of this attempt see: Teece et al. (1994) and Neffke and Henning (2013).

certain goods is prerequisite for the production of others. To do this the authors calculate the increase in the statistical probability that a good (a) will be produced based on the presence of good (b). The resulting matrix is the taxonomy matrix.

In addition to the taxonomy matrix, Zaccaria et al. (2014) introduce the assist matrix, which quantifies the number of times the production of a good (b) in time (t) leads to emergence of good (a) in time (t+$\delta$)[3].

While the taxonomy matrix is estimated for a single year of trade data, the assist matrix is estimated by comparing multiple periods.

A similar approach to the assist matrix was made by Quan et al. (2017). These authors created a directed network for which the links are the conditional probability of exporting good (a) given that a country exports good (b) in (t) years. In this paper they also separately estimate the probability of going from (a) to (b) and from (b) to (a).

Another important contribution is from the work of Pugliese et al. (2019), who scrutinized the diversification process within the technology network by creating a directed space using a one-step process. They built their network based on patents. For every couple of patent classes, they counted the number of times patent class (a) had become active in a region (r) that had already been active, for some time, in patent class (b). This sum was normalized by the ubiquity level of patent class (a)[4] and the level of diversification[5]. ($\delta$) corresponds to the time lag. An important contribution of this work was to assess the significance of links by comparing the probability that a link exists with a multiple random network (null network) of similar parameters (patent class ubiquity and regional diversification). The multiple random networks are generated from a null model based on the bipartite configuration model from Saracco et al. (2015). This creates a robust directed space that can help better understand the real interdependence between technologies. More specifically, instead of takingonly the strongest links, which is commonly done in the literature, it will detect links that are less strong but statistically significant.

Further enriching this body of literature, Napolitano et al. (2018) delved into the autocatalytic structure of the network by investigating the dynamics of autocatalytic aets (ACSs) over time. An ACS is a subset of nodes within the directed network such that each node receives at least one catalytic link from other members of the set. The ACS framework introduced by Jain and Krishna (1998) and formalized in Jain and Krishna (2003) describes the process by which complex network structures emerge and evolve in ecosystem networks. The authors discerned varying dynamics within the catalytic network, noting that the core of the ACS demonstrated a higher growth than nodes at the periphery or outside the set. The authors also noted the expanding nature of the ACS over time, suggesting that research intensity in the ACS

[3]$\delta$ is a time difference introduced to allow for the creation of new production capabilities and therefore the exportation of new goods.

[4]The number of regions active in patent class (a).

[5]the number of patent classes active in region (r).

can influence the distribution of links and increase the probability of the emergence of a new link towards a previously unrelated node. A similar process was detected in the patent network in Napolitano et al. (2018). They found that patent classes in the core had higher growth. They also found that the nodes in the periphery and outside of the set had a greater relianceon core technologies, namely *computer science*.

Closer to our interest, Cunzo et al. (2022) use an approach similar in spirit to the directed patent spaces in Napolitano et al. (2018) and Pugliese et al. (2019) to investigate the empirical relationship between green technologies and industrial production. The study used patent data as a proxy for green innovation and export data as a proxy for industrial production. The authors constructed a bipartite directed network to track the influence of green technological development on industrial production, based on time-lagged co-exportations of green technologies and products in countries' industrial and technological specialization profiles. The network links were then refined using a maximum entropy null model. The study found a strong connection between green technologies and the export of products associated with the processing of raw materials, crucial for climate change mitigation and adaptation technologies. The results also suggest the increasing importance of more complex green technologies and high-tech products in the network over time, indicating the growing significance of these elements in green economic growth. This work contributes to our understanding of how green technology innovation is integrated into industrial production and evolves over time.

We also link this literature to another body of literature which is the green complexity literature, introduced in Hamwey et al. (2013). This contribution proposes a green product space methodology to identify export strengths of countries in green markets. The model identifies green products for which a country is likely to be globally competitive based on the performance of related products. This method, illustrated with the case study of Brazil, provides a fundamental tool to support "green growth" and sustainable development.

The goal of such models is to use the economic complexity approach in order to understand the process of path dependency in a country's specialization and its effect on the adoption of green goods (Talebzadehhosseini et al., 2019; Talebzadehhosseini and Garibay, 2021). This research also helps to inform green economic policy-makers about which green industries are easier to establish (Fraccascia et al., 2018; Barbieri et al., 2022; Alhaddad et al., 2023; Müller and Eichhammer, 2023). Ferraz et al. (2021) provides a broader review of the literature, identifying research gaps in integrating environmental and social sustainability considerations in economic diversification and complexity studies. The authors call for more cross-fertilization between economic complexity research, "green growth" policy research, and studies on diversification in less developed regions. Our main contribution to the literature will be to study the positive feedback loops between green and non-green economic activities through the process of industrialization.

# 3 Data and Method

## 3.1 Data

In our analysis, export data are used to construct a proxy of countries' production capabilities, a method widely followed in the literature, as explained in the previous section. Our database is extracted from the *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII). The *Base pour l'Analyse du Commerce International* (BACI) Dataset, referenced by Gaulier et al. (2008), is a comprehensive data set focused on international trade. It encompasses an extensive catalog of over 5000 products, each denoted with a unique 6-digit identifier. This data set meticulously logs the annual aggregated value[6] of product flow from one country to another, thus providing a comprehensive overview of international trade patterns. Moreover, it also includes detailed information about both the country of origin and the destination for all bilateral transactions, thereby covering a global spectrum of approximately 200 countries.

Another advantage of this database is that it is based on the Harmonized System (HS). The HS is a standardized product nomenclature used by customs authorities and managed by the World Customs Organization, which provides a regularly updated list of traded goods. The exporter's and importer's reports of the same flow are linked and the reliability of each reporter is assessed. Based on this assessment, an accurate estimation of countries export value has been computed by the CEPII. We use the version of the database that relies on the HS 92 nomenclature which covers every year between 1995 and 2021.

## 3.2 Green goods classification

Classifying green goods is a complex task. There is no international consensus on which goods are considered as environmental. To create a list of such goods, many studies have been conducted by working groups of the Organisation for Economic Co-operation and Development (Sauvage, 2014, OECD) as well as work carried out during negotiations for trade and tariff purposes by the World Trade Organization (WTO, 2011) and the Asia-Pacific Economic Cooperation (APEC, 2012). In this paper we use the Combined List of Environmental Goods (CLEG) used by the OECD and introduced in Sauvage (2014). This list has been chosen for multiple reasons: It is independent of trade negotiations and their mercantilist influences (Balineau and De Melo, 2013; de Melo and Solleder, 2017). Negotiated environmental product lists, such as the WTO list, exclude some goods whose environmental content is obvious and include other goods for which the environmental aspect may be debatable. The CLEG contains 248 environmental goods at the 6-digit level, following the HS 07 nomenclature, and can be divided further according to their environmental aspect.

[6] The value is estimated in current USD.

The CLEG list goes through several checks to ensure its environmental content is accurate, including insights from Environmental Business International Inc. (EBI)[7] and from the OECD.

The CLEG list's main downside is the multiple end-use of some of its goods. These goods can therefore also be used for non-environmental purposes[8]. The categories of environmental goods in the CLEG database are the following:

- Air pollution control
- Cleaner or more resource-efficient technologies and products
- Environmentally preferable products based on end use or disposal characteristics
- Heat and energy management
- Environmental monitoring, analysis and assessment equipment
- Natural resource protection
- Noise and vibration abatement
- Renewable energy plant
- Management of solid and hazardous waste and recycling systems
- Clean up or remediation of soil and water
- Waste water management and potable water treatment

Using publicly available correspondence tables[9], we transformed the data from HS 2007 to HS 1992 to ensure compatibility with the trade data. It is worth noting that no goods were omitted from the original classification, but a total of 24 goods had changes in their assigned HS codes.

### 3.3 Building the Directed Product Space

We construct a directed network of products by taking inspiration from the method used by Napolitano et al. (2018) in their study of a directed network of technological fields. In our framework, the nodes of the network are the products and the network edges are 'catalyzing' relationships: a product is connected to another if the presence of the first product in a country's export basket at time (t) increases the probability of the emergence of the second good in the same country's export basket at time (t) + $\delta$.

[7]Environmental Business International Inc. (EBI) is a publishing and research company that generates strategic market intelligence on emerging opportunities in the Environmental Industry and the Climate Change Industry (Environmental Business International inc., 2023).

[8]Green goods can be classed in two categories. Goods for Environmental Management (GEM) are defined as products (and services) that reduce environmental risk and minimize pollution and resource use. These goods are generally considered as multiple-end use, which includes pollution management and resource management. Environmentally Preferable Products (EPPs) are those that produce less environmental damage either in production, use, or disposal. These goods are generally considered as single-end use.

[9]The correspondence tables between HS 2007 to HS 1992 can be found at: https://unstats.un.org/unsd/classifications/Econ.

An implicit assumption is made using trade data: a good is exported only if that country masters the production capabilities of that good. However, in order to filter out the countries that export only a marginal quantity of a good, and thereby avoid a situation in which every countries has production capabilities for every good, research in the product space (also that in the Industry space and the Technology space) relies heavily on the concept of revealed comparative advantage, introduced by Bena Balassa (Balassa, 1965).

$$RCA_{\mathrm{cp}} = \frac{(X_{\mathrm{cp}}/\sum_{\mathrm{p}} X_{\mathrm{cp}})}{(\sum_{\mathrm{c}} X_{\mathrm{cp}}/\sum_{\mathrm{c}}\sum_{\mathrm{p}} X_{\mathrm{cp}})} \tag{1}$$

$X_{\mathrm{cp}}$ corresponds to the value of the export of the good (p) by the country (c). The RCA index measures the share of the export of a given good in the country's export basket ($X_{\mathrm{cp}}/\sum_{\mathrm{p}} X_{\mathrm{cp}}$) compared to the world share of the same good over the world total export ($\sum_{\mathrm{c}} X_{\mathrm{cp}}/\sum_{\mathrm{c}}\sum_{\mathrm{p}} X_{\mathrm{cp}}$). A value of RCA $> 1$ implies the country is better at exporting a good than most other countries.

The fundamental concept of comparative advantage suggests that a country has an interest in specializing in a narrow subset of goods, even though it could potentially produce a broader range of goods. In contrast, the Capability Theory explains why countries specialize far less than is predicted by the comparative advantage theory. Countries are expected to export, or at least produce, all the goods for which they have the required capabilities. It is therefore important to understand that the RCA index captures an aspect of specialization not present in Hausmann's comparative advantage theory.

This approach has some conceptual limitations. It is not able to discriminate between instances where certain countries are unable to produce a specific product because they lack capabilities and instances where countries fail to reach a significant level of export for these goods due to market forces or a lack of competitiveness. In our methodology, we use the RCA concept not to assess the competitive advantage of countries, but rather to assess the presence of production capabilities.[10]

Eq. 1 enables us to construct the Country-Product Matrix ($\mathrm{M_{cp}}$), which provides the level of competitiveness for all exports of every country. To model the catalytic links between all the exports for which the country is competitive, we model the export basket as a directed product space. The adjacency matrix of our network is defined as follows:

$$M_{\mathrm{cp}} = \begin{cases} 1 \text{ if RCA }_{\mathrm{cp}} > 1 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

This is a binary matrix that has a value of 1 if a country has a RCA>1 in the exportation of a given good and 0 otherwise. The 'ubiquity' of a product is defined

[10]Our next step in the publication of this research will be to confirm our results by doing a robustness check with a lower RCA threshold.

as follows:

$$u_{\mathrm{p}} = \sum\nolimits_{\mathrm{c}} M_{\mathrm{cp}} \tag{3}$$

This corresponds to the number of countries that export the good that have RCA>1. Likewise Eq. 4 defines the diversification of country (c), which is the number of goods exported by that country.

$$d_{\mathrm{c}} = \sum\nolimits_{\mathrm{p}} M_{\mathrm{cp}} \tag{4}$$

Finally, we define the probability that a product p favors the emergence of a product p' as follows:

$$B_{\mathrm{pp'}}(t, t+\delta) = \frac{1}{u_{\mathrm{p}}(t)} \sum_{c} \frac{M_{\mathrm{cp}}(t) M_{\mathrm{cp'}}(t+\delta)}{d_{\mathrm{c}}(t+\delta)} \tag{5}$$

This probability measure consists of the product between the Mcp at time (t) and the Mcp matrix at time $t+\delta$.[11] Each matrixis normalized by the ubiquity of each product ($u_p$) and the diversification of each country ($d_c$).

Eq. 5 is a directed product space. The undirected product space described in Pugliese et al. (2019), has similar characteristics to a directed product space, which is the network representation of the stronger probability of exporting the good (a) $\delta$ years later if the good (b) is exported by a country (c) in year (t).

## 3.4 Assessing Link Significance

The adjacency matrix defined above expresses the probability of catalysis between all pairs of goods. A high value for the catalyzation of product (b) by product (a) means that if the product (a) is in the export basket of a country at time (t) it is highly likely that product (b) will be found in the export basket at time (t+$\delta$). Likewise, a probability of 0 implies that when product (a) is exported at time (t) product (b) is never found in the export basket at time (t+$\delta$).

The next step is to discriminate between the instances in which a product catalyzes another and the instances in which both products coincidentally happen to be produced by the same type of countries. We therefore filter out the links for which a co-exportation may be coincidental. A high level of co-exportation between goods can occur for three reasons: (i) Some products have a high co-exportation rate because they are exported by most countries (high ubiquity). (ii) Some products have a high co-exportation rate because they are produced by a few highly industrialized and diversified countries. (iii) Some products have a high co-exportation because they are based on the same capabilities or because they are complementary.

[11] The result is a product by product matrix which corresponds to a inter-temporal matrix of co-exportation. In this matrix, the higher the value of an entry, the higher the amount of co-exportation.

To select only links that are economically relevant (and correspond to (iii)), we follow the methodology presented in Napolitano et al. (2018) and Pugliese et al. (2019). For each link we measure the likelihood it emerged coincidentally. This is done by comparing the co-exportation probability of a product at time (t) and another product at time (t+$\delta$) in the empirical network with the co-exportation probability of the same products in 1000 random networks. If the co-exportation of the two goods occurs a significant number of times in the random networks then the catalyzing link is considered as economically non-relevant. More precisely, we assess the significance of each link depending on the frequency of its occurrence in 1000 directed product spaces with randomized $\mathrm{M_{cp}}(\mathrm{t})$ and $\mathrm{M_{cp}}(\mathrm{t}+\delta)$ matrices. The $\mathrm{M_{cp}}$ matrices are generated using a null model, based on the bipartite configuration model presented in Saracco et al. (2015) that can generate a randomized binary matrix with specific degree distribution. It is used to create binary matrices in which countries have on average the same diversity ($\mathrm{d_c}$) and products have on average the same ubiquity ($\mathrm{u_p}$) as in the original $\mathrm{M_{cp}}$ matrices. A first null model is used to randomly generate 1000 binary matrices that display the same average degree distribution as the original $\mathrm{M_{cp}}(\mathrm{t})$. Another null model is used to generate 1000 null matrices for $\mathrm{M_{cp}}(\mathrm{t} + \delta)$. Those two sets of null matrices are used to create 1000 randomized directed product spaces (null network).

The following step is to test the empirical directed product space against the 1000 randomly generated counterparts, and to select a threshold for the number of occurrences over which the link will be discarded. This threshold will be our p-value. Only links that are statistically significant depending on the selected p-value are considered as economically relevant.

In this paper, we have selected a significance level of 0.1%. Individually comparing each link's p-value to our significance level creates an issue called the multiple comparison problem. This problem arises when an analysis involves multiple statistical tests. Running such a high number of tests increases the probability of making false positive errors Colquhoun (2014). To accurately control for the proportion of false positives simultaneously for all links, we adjust the confidence level. This is done by applying the false-discovery-rate correction procedure from Benjamini and Hochberg (1995) to maintain the overall significance level at 0.1%.

Finally, we convert our filtered network into a binary format, where a link represents a significant catalyzation relationship, and the absence of a link indicates no significant catalyzation connection.

## 3.5 Detecting Autocatalytic sets in the network

We apply the ACS framework to understand the multiple feedback loop effects that occur within the product space.

The application of the ACS framework on the directed product space will provide new information on the positive feedback loop effect between the different exports of a country. This is especially relevant to understand the interdependences between green goods and exports that can be negative for the environment.

Our detection of ACSs utilizes a methodology detailed in Jain and Krishna (2003) and Napolitano et al. (2018), based on the Perron-Frobenius theorem. This approach involves representing the directed product space as a binary adjacency matrix, indicating the presence (1) or absence (0) of a link. This binary structure makes the adjacency matrix non-negative.

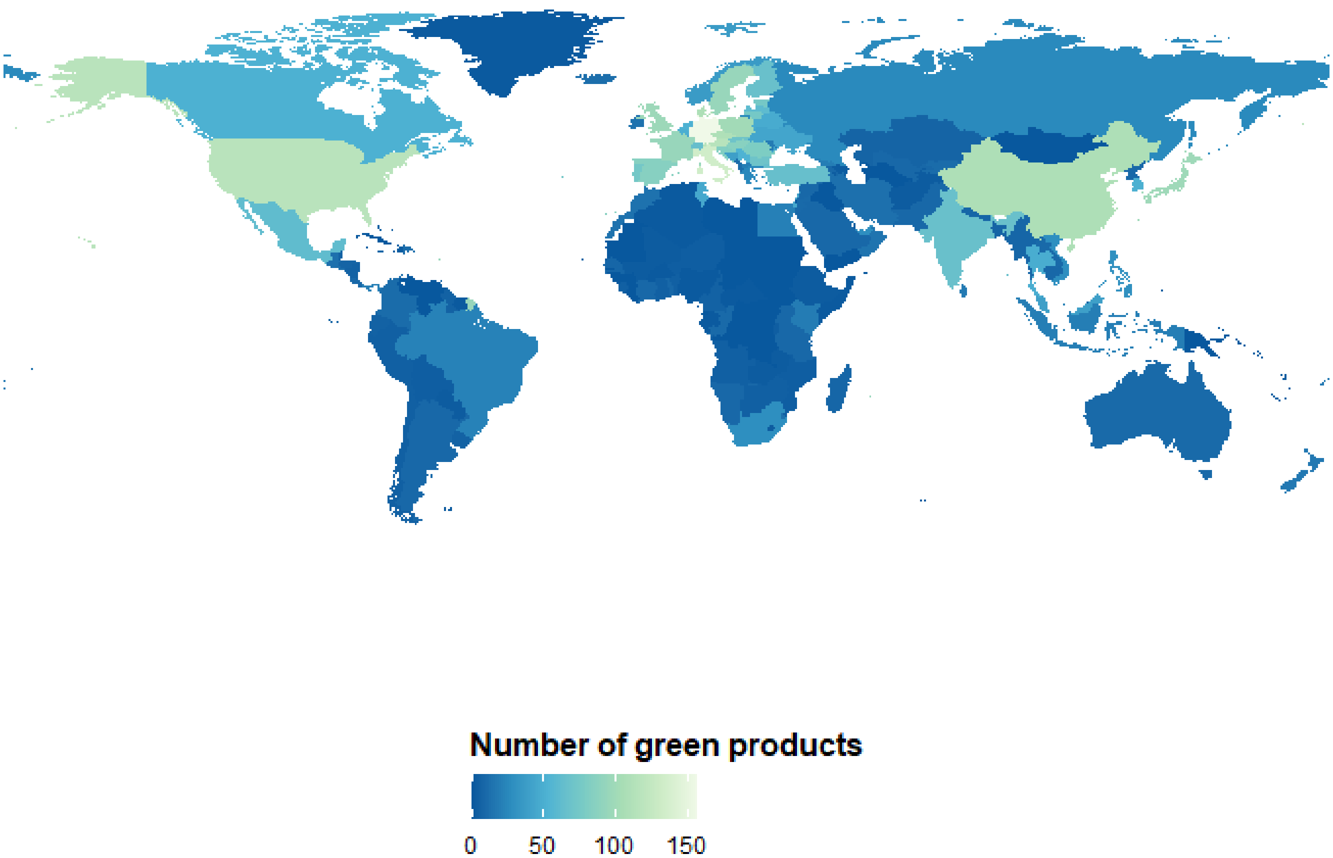
The Perron-Frobenius theorem guarantees that for all non-negative matrices there exist a real positive eigenvalue that is the largest of all the eigenvalues. This eigenvalue, the Perron-Frobenius eigenvalue (PFe), signifies the presence of closed directed paths or feedback loops within the graph. When the PFe exceeds 1, we have an ACS, indicating that certain goods are part of self-reinforcing sets, benefiting from these feedback loops. The PFe value also suggests the growth speed due to the feedback: the larger the PFe, the faster the anticipated growth.

Nodes forming part of an ACS can be identified through the corresponding eigenvector of the matrix. If an eigenvector element is non-zero, its related node is part of an ACS.

# 4 Results

We start by highlighting some stylized facts about which countries produce green goods and at which stage of development. Figure 1 shows the number of green products exported (with RCA > 1) per country.

Figure 1: World map of diversification level in green products




*Note: Areas with lighter green are characterized by higher green product diversification*

The map shows that the production of green goods is not evenly distributed. Green goods are mainly produced by developed countries in North America, Europe and East Asia. This geographical representation underscores a disparity, revealing that emerging economies are laggard in the export of green products.

This is reinforced by the following two figures. First, we plot in Figure 2 the number of green goods exported by a country against its level of diversification. Figure 3 compares the number of green goods exported by a country against a measure of the country's industrial sophistication. This index, based on the work of Tacchella et al. (2013), is a sophistication indicator for both products and countries. Each product is given a value based on its complexity level, which is essentially determined by the complexity of the capabilities required to produce that good. Countries are assigned a value according to their ability to export products that necessitate a complex combination of capabilities. For more information, see A.2.2.

Figure 2: Green goods diversification index versus total product diversification index

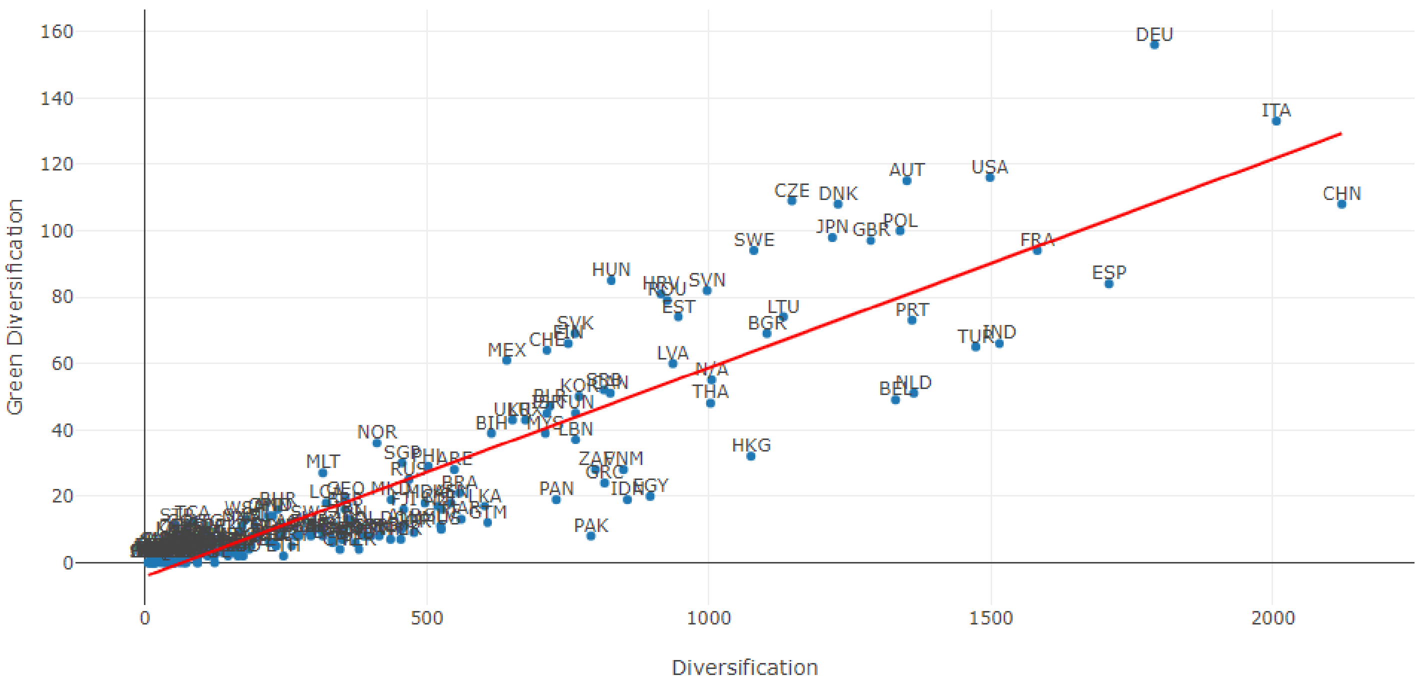
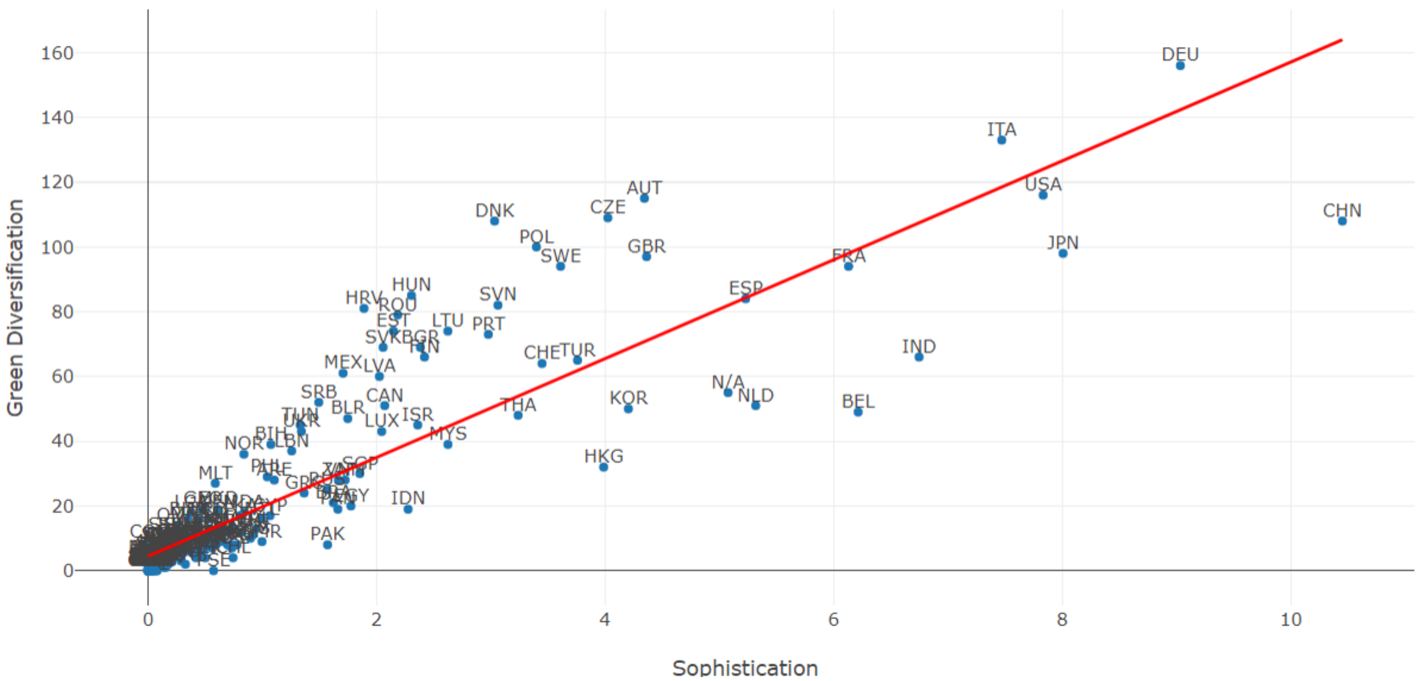


Figure 3: Green goods diversification index versus country sophistication index




Both graphs show a positive linear correlation between the level of green diversification and the level of industrial development. From an economic perspective, this suggests that production capacity and market dynamics of environmentally sustainable goods can be a by-product of a country's economic progress. The acquisition of capabilities in the production of green goods seems to be obtained in a similar way as for other goods in the course of industrialization. Countries with a stronger capacity to produce complex products will be able to access more complex green goods categories. Leaders in the production of green goods are also those capable of producing other goods, such as Germany, Italy, China, and the United States, showing the complementary in the development of production capabilities.

The rest of this section is organized into two distinct parts. The first delves into an analysis of the network's overall structure, and the second zooms in on the evolution of the ACS within the network and its ramifications for industrial development.

In the first part, we examine the position of green goods within the overarching structure of the product space, including their position in the ACS. We explore the probability that green goods are adopted contingent on their placement in the product space and the potential influence this might have on future diversification. Furthermore, we also investigate the various feedback loops identifiable within the directed product space.

In the second part, we focus on how multiple overlapping feedback loops can impact a country's environmental policy.

## 4.1 Analysis of the Directed Product Space

The directed product space for the years 2017-2019 is displayed in Figure 4. Each node represents a product at the 6-digit level, while links express a significant likelihood of exporting a good (b) two years after good (a) is exported. In this chapter, the results refer to a forward period of two years ($\delta = 2$). A thorough network analysis of the product space over the years can be found in A.3. This figure shows the main subgraph of the network, which corresponds to 2641 of the 4647 products in the database. All the other products are isolated nodes, which are not displayed in the figure.

Figure 4 displays two interconnected subgraphs. The smaller subgraph on the left consists entirely of textile products and lacks any green goods for the period 2017-2019. A second, more diverse subgraph is visible on the right, featuring a large variety of industries.

Out of the 4647 products listed in BACI for the years 2017-2019, 2641 belong to the directed product space, accounting for 56.8% of the total. Within this directed product space, products that are ACS members make up 1886 items (or 40.5% of the total products), while non-ACS members account for 755 items (or 16.2%). Among the ACS members, 1028 are identified as core products (22.1%), and 858 are identified as peripheral products (18.4%).

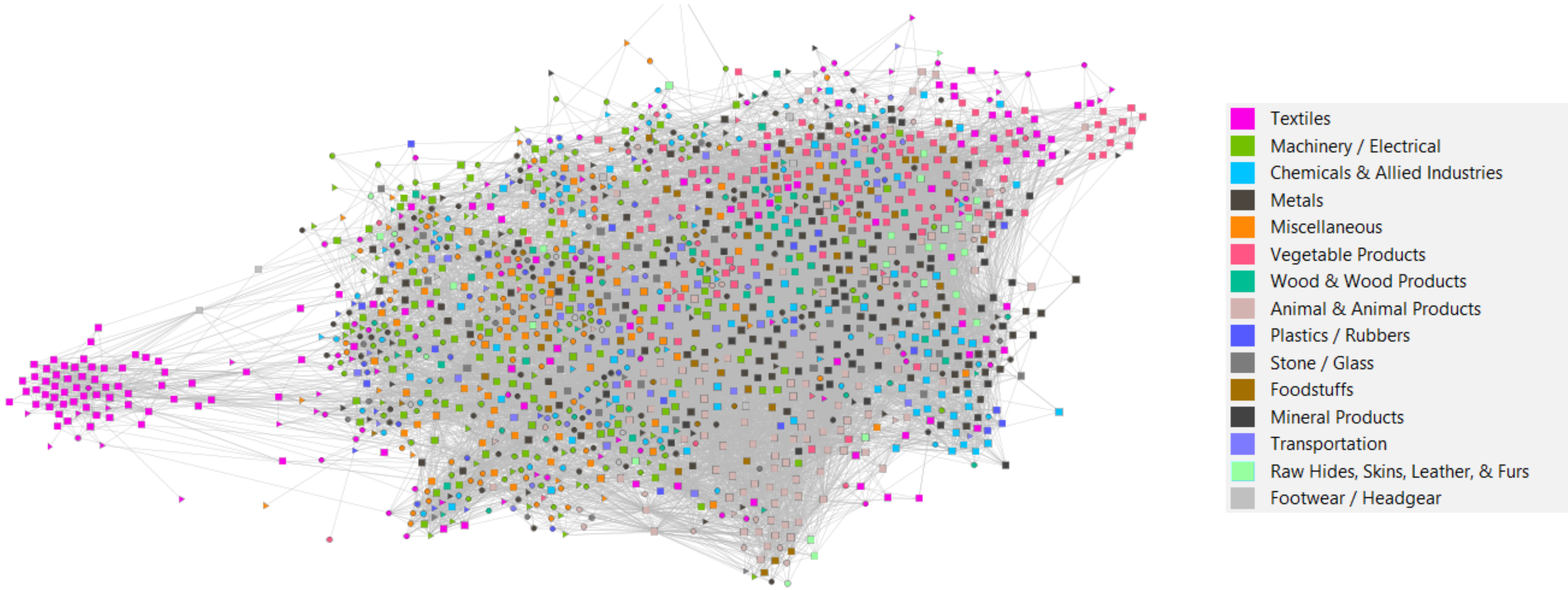
The observation that a majority of nodes within the directed product space are also part of the ACS is not inconsequential. This suggests that the structure of the directed product space is influenced by the ACS framework. The ACS is defined as a set of nodes where each member receives at least one link from another member within the set. Core members are specifically those ACS members that send a link to other core members.

Nodes not included in the directed product space are generally have no links or only a few links. This group of 2006 products represents less than half the total. Only the nodes represented organize themselves into clusters. Most nodes receive at least one link from other ACS members, and the majority of ACS members are classified as core. Figure 5 illustrates the ACS within the directed product space.

Another observation concerns the distribution of green goods within the main subgraph. In this context, green goods are distributed homogeneously across the directed product space, with the exception of the textile industry. This pattern is attributable to the nature of the CLEG classification for green goods, which is based on their potential use in sustainable transition processes rather than any inherent product quality. More precisely, green goods are not defined by a specific set of capabilities or technologies. Instead, they encompass a diverse range of specific applications, some of which may contribute to sustainable development.

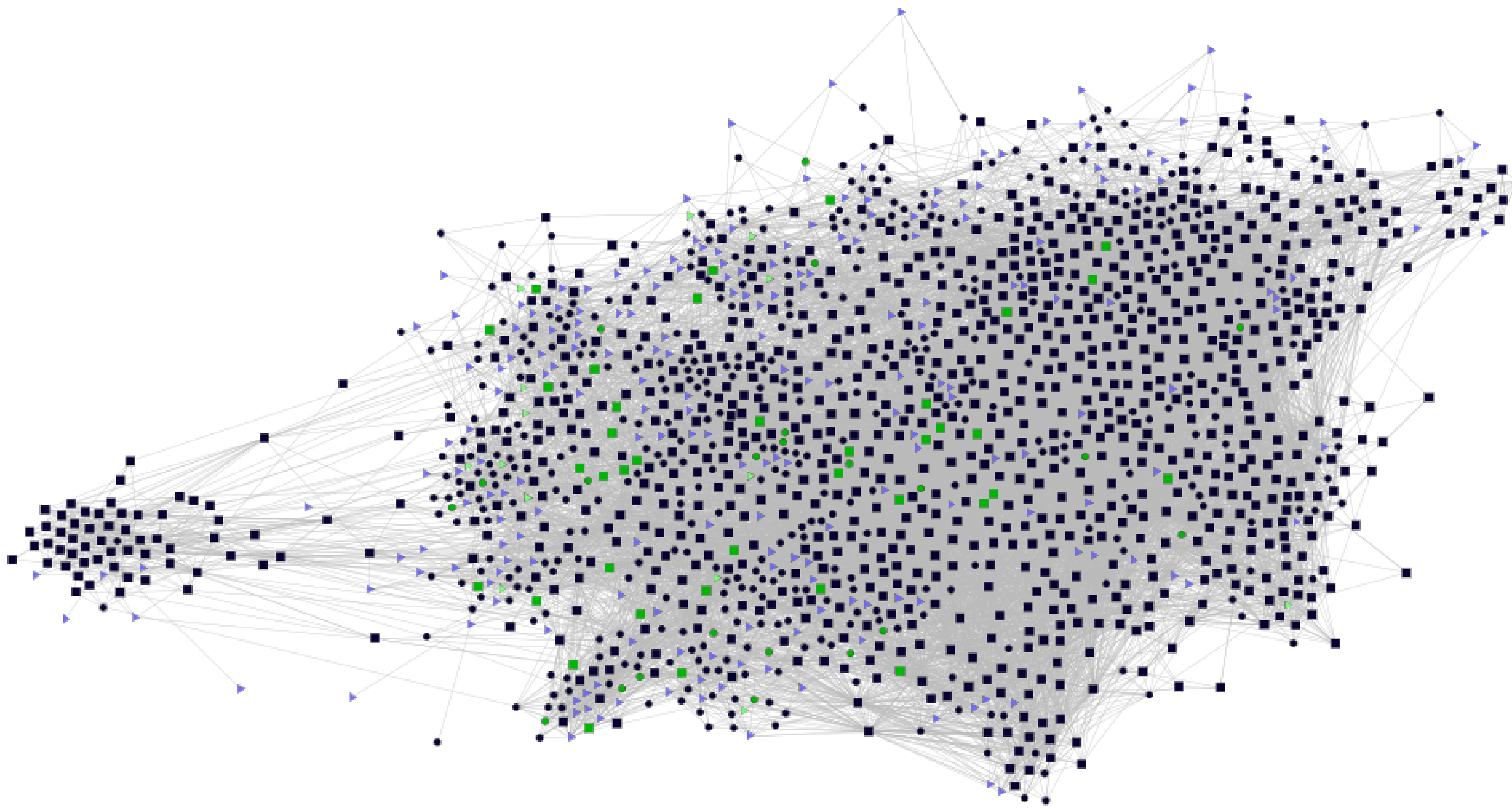
This has multiple implications for policy-making. First, the absence of specific

Figure 4: The directed product space for the period 2017-2019




*Note: The figure represents the directed product space for the period 2017-2019. Each node corresponds to a specific exported good, while the links between nodes indicate catalyzation relationships between the goods. The colors are used to differentiate various sectors of the economy. Only nodes that are connected through links are displayed in the visualization. The network visualization was generated using the ForceAtlas2 force-directed algorithm (Jacomy et al., 2014). For better comprehension, it is recommended to view this figure in color.*

Figure 5: Green goods and the autocatalytic set in the directed product space for the period 2017-2019.



*Note: The figure illustrates the positioning of green goods and the ACS within the directed product space for the period 2017-2019. Green nodes represent Green goods (Dark green corresponds to ACS members and light green corresponds to non-ACS members). Black nodes are non-green ACS members. Blue nodes are non-green non-ACS nodes. The shape of the node conveys information about the position in the ACS: square nodes are the core, round nodes are the periphery, and triangle nodes are outside of the ACS. For better comprehension, it is advisable to view this figure in color. Only nodes that are connected are displayed in the visualization.*

clusters where green goods are concentrated suggests that there is no singular set of capabilities that can optimize the path toward producing these goods. As a result, the development of green goods will require the acquisition of a diverse set of capabilities, each tailored to a specific green good. Second, the products most similar to green goods in terms of capability requirements are often non-green goods. This suggests that gaining a comparative advantage in the production and export of non-green goods could be a useful stepping stone toward acquiring the capabilities needed for green goods. Lastly, countries may find it beneficial to prioritize the development of green goods that are closely related to products they already export, thereby leveraging existing capabilities for sustainable development.

This creates a direct connection between industrial development and the export of green products. Enhancing industrialization levels in developing countries is essential for increasing the number of nations capable of producing green goods. In this scenario, the concept of leapfrogging becomes particularly relevant. It offers these countries the opportunity to skip certain phases of industrial development, enabling them to adopt green technologies more quickly (leapfrogging).

The directed product space serves as a roadmap in this context, indicating which products can be leapfrogged and which cannot, based on a country's existing export portfolio. This approach could accelerate the transition to sustainable practices and simultaneously boost export capabilities in green goods.

## 4.2 Green and non-green interactions in the Network

In this subsection, we delve into the intricate interplay between green and non-green products in the network, with a particular focus on their catalyzation relationships. Table 1 displays a summary of the links between green and non-green goods in the network. It quantifies the pairwise catalyzation relationship between green and non-green.

Table 1: Contingency table summarizing the type of good at the origin and destination of a link.

| | Number of links | | | % of incoming links | | | % of outgoing links | | |
|---|---|---|---|---|---|---|---|---|---|
| From/To | Green | Non-green | Total | Green | Non-green | Total | Green | Non-green | Total |
| Green | 46 | 630 | 676 | 6.8% | 93.2% | 100% | 2.95% | 5.24% | 3.04% |
| Non-green | 832 | 20702 | 21534 | 3.86% | 96.14% | 100% | 97.05% | 94.76% | 96.96% |
| Total | 878 | 21332 | 22210 | 3.95% | 96.05% | 100% | 100% | 100% | 100% |

*Note: period 2017-2019. The number of links is the number of catalytic relationships from one category to another. % of incoming links measures the share of green and non-green products that catalyze each given category. % of outgoing links measures the share of green and non-green products catalyzed by each category.*

From Table 1, it can be inferred that in the directed product space for the period

2017-2019, green goods are more likely to act as catalysts for non-green goods, as evidenced by 93% of outgoing links from green goods being directed towards non-green products. Intriguingly, this statistic is marginally lower than its non-green counterparts. Similarly, green goods are also more likely to be catalyzed by non-green goods. It is observed that an overwhelming 97.05% of incoming links going towards green goods originate from non-green goods.

What the data essentially suggests is a complex interdependence between green and non-green goods, where non-green goods appear to be instrumental in fostering the growth of green industries.

To better understand the complex choices between different types of goods, we have provided a comparison of the centrality of green goods and non-green goods in the appendix (see Appendx A.4). This comparison shows that non-green goods have a higher centrality than green goods.

It also shows that green goods receive fewer catalytic effects than other products. This reinforces the idea that specific industrial policies are necessary to orient the industrialization process towards green products, even when the capabilities are present.

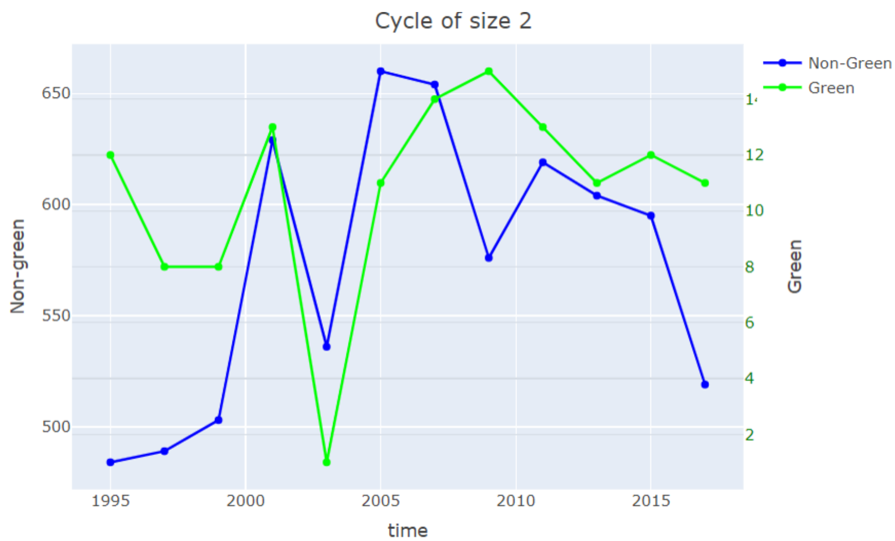
We will now focus on the interdependence between the green and the non-green goods. More specifically, we will investigate positive feedback loops that occur within the product space between the exports of those two types of products. Consider, for instance, the feedback loop between product 840810 (Engines: for marine propulsion, compression-ignition internal combustion piston engines (*diesel or semi-diesel engines*) and product 560811 (*Twine, cordage or rope: fishing nets, made up, of man made textile materials*). The former is a non-green good, an engine used in motorboats and the latter is a green good, a twine that has a higher rate of recyclability, durability, reduced environmental impact and responsible disposal than other twines.

While this study makes no assumptions about the reasons for catalysis, it is reasonable to deduce in this case that a nation with a robust shipbuilding industry would likely support both an engine and a twine industry. Consequently, the growth of one industry could potentially stimulate the growth of the other, forming a mutually beneficial loop.

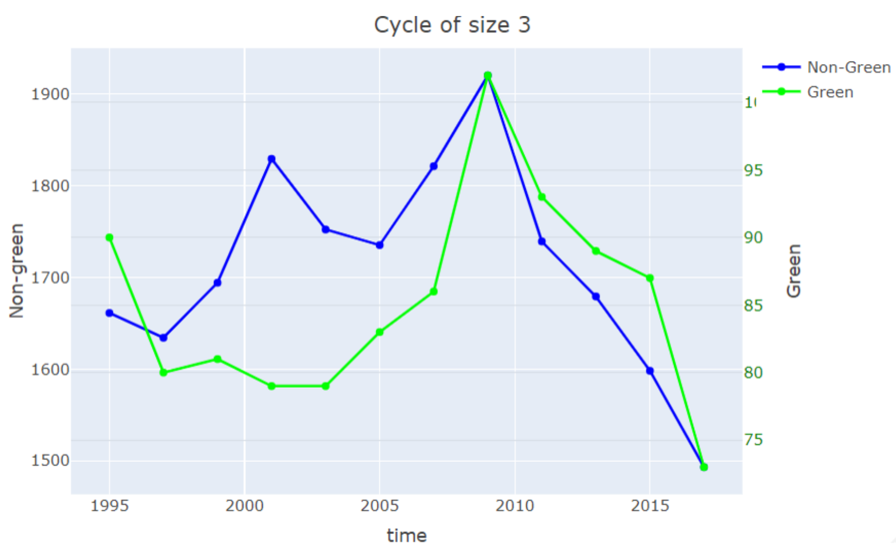
In the following, we generalize the study of feedback loops and their effect on the catalyzation of green products.

We examine the autocatalytic cycles in Figures 6a and 6b, which capture cases where green and non-green goods are part of a feedback loop within the network. It is interesting to note that only a small fraction of green products are part of these cycles with other green goods. This could be ascribed to the diverse nature of manufacturing processes, and uses associated with the green goods.

Figure 6: Membership of green and non-green goods in two- and three-node cycles in the directed product space




(a) Two-node cycles




(b) Three-node cycles

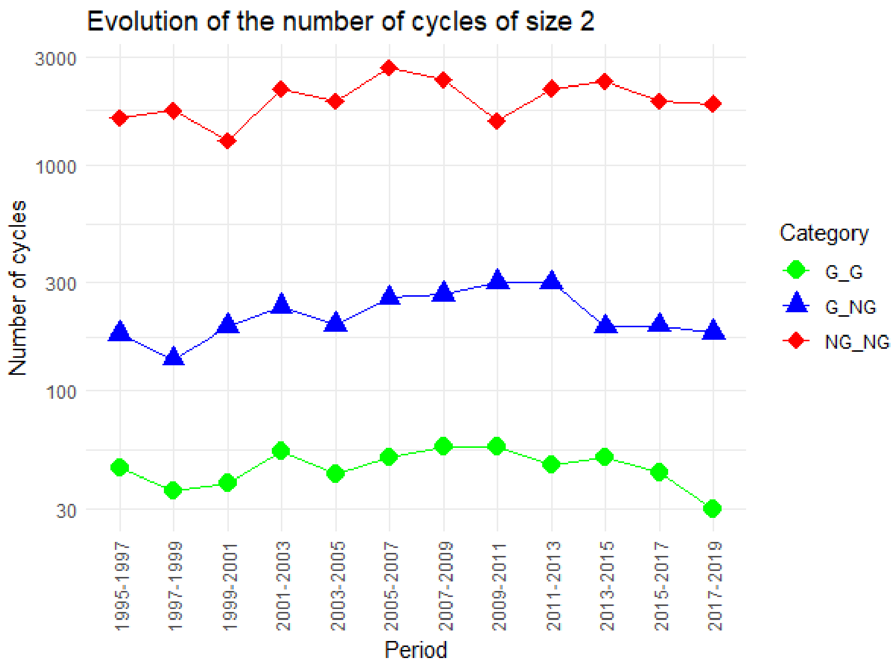
Given the broad spectrum of technologies they originate from, green goods do not necessarily share common capability requirements. However, they are inclined to form connections with other product types within the same industry or those based on similar technologies, most of which are non-green goods.

Figure 7a illustrates the evolution of the number of two-node cycles that exist between green goods, non-green goods, or between one green and one non-green good. The number of autocatalytic cycles containing only green goods is notably low. This emphasizes the concept that a significant portion of green goods' growth may be catalyzed through interactions between green and non-green goods. Two-node cycles involving green goods are predominantly between a green and a non-green good.

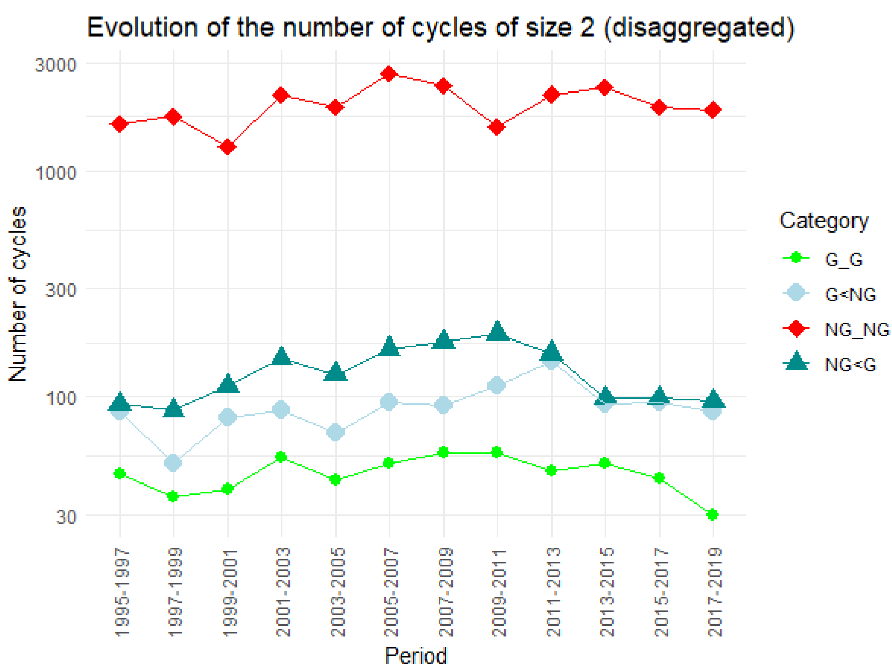
Subsequently, in Figure 7b, we explore the PRODY Index (see Appendix A.2.3 for the detail of the construction of this index) of the products within each cycle. In essence, we are examining which product is likely to emerge first in the development process. The results reveal that for every year between 1997 and 2011, in two-node cycles involving both green and non-green goods, the non-green good is more likely to be the first to emerge and hence catalyze the growth of the green good. However, in the years succeeding this period the numbers of two-node cycles converge. One limitation of these counts is that the cycles are often not standalone, but rather frequently overlap with each other. Some products are part of multiple cycles, resulting in a cumulative effect of these cycles. To address this and account for such cumulative cycles, the next subsection will implement an ACS approach.

Figure 7: Number of two-node cycles by product type

(a) Number of two-node cycles




(b) Which of the G & NG has a lower prody




*Note: PRODY is an indicator corresponding to the weighted average of the GDP per capita of the countries that export a given good. In this context it is used as a proxy to measure at which level of development a good emerges in an export basket. "G<NG" means that the green good has a lower prody that the non-green good. "NG<G" means that the non-green good have a lower prody that the green good. The source for the GDP per capita is CEPII (2023)*

## 4.3 Autocatalytic sets, green goods and product space

### 4.3.1 The structure of the ACS

In this section, we delve deeper into the analysis of product cycles by employing the ACS approach, which allows us to account for overlapping cycles. The economic implications of the ACS in the product space are important for economic development.

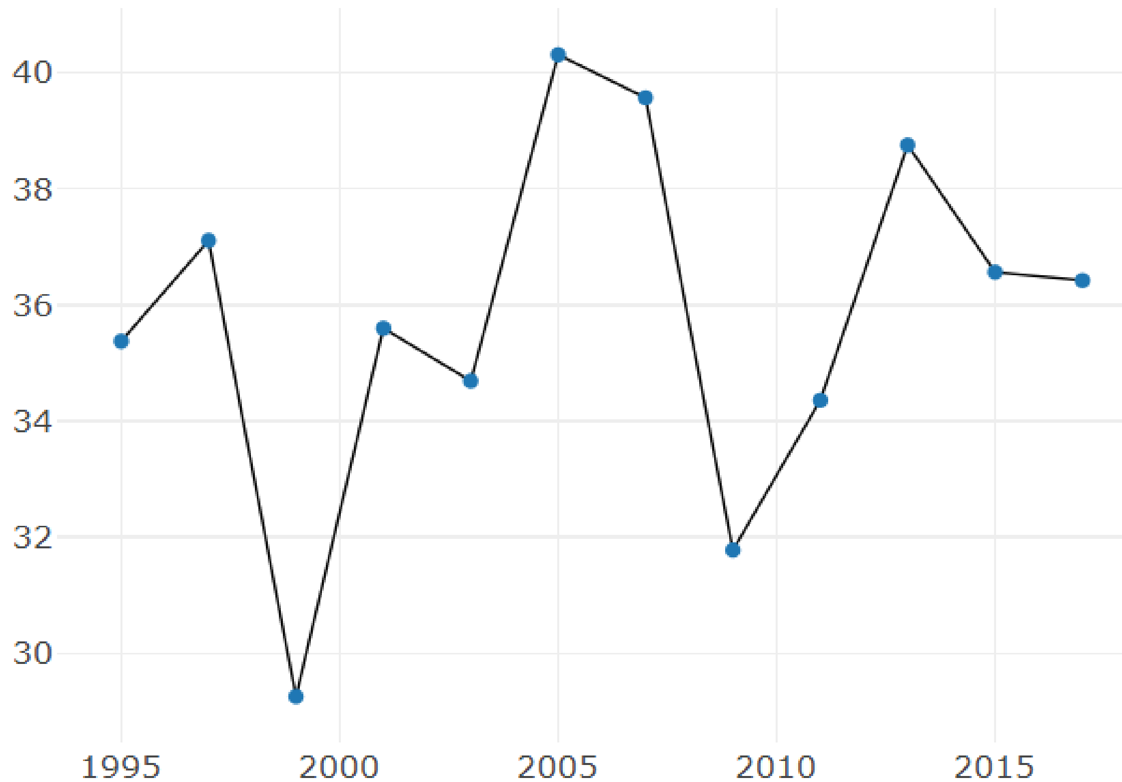
The ACS encompasses all the products that can be catalyzed by other members of the set. Products that are ACS members are more likely to benefit from a boost in their exports if the country has a comparative advantage in the exports of other products. Exporting a product from the ACS, and especially from the core, implies the ability to leverage its structure in order to guide the country's diversification process. The more goods from the set that are produced, the stronger the effects will be from the multiple overlapping feedback loops, and the more potent the ACS catalyzation effects will become.

It is therefore crucial for a developing country, which lacks resources, to leverage the ACS to benefit from a boost in its industrialization. Using the ACS structure to develop green industries is also feasible. The structure of the network can provide multiple catalyzing effects for the development of green products. As a country increases the number of products in which it has a comparative advantage within the set, it becomes easier for that country to develop green products. Conversely, as a country increases the number of green products in the set, it is more likely to adopt other non-green products within the set. Understanding the position of green products within the set is essential for comprehending the link between diversification and the adoption of green products in developing countries.

We will first describe the ACS in the product space and then consider the role of green products within this set.

One way to measure the autocatalytic feedback loop present within the network is through the strongest eigenvalue of the network, known as the Perron-Frobenius eigenvalue (PFe). An eigenvalue greater than 1 indicates the existence of an ACS. The magnitude of this eigenvalue reflects the autocatalytic strength of the network – the larger the value, the stronger the autocatalytic effect. As depicted in Figure 8, the PFe in our network fluctuates over time. It decreases from 1995 to 1999, rises till 2005, drops again until 2010, and finally increases, reaching 36 in 2017.

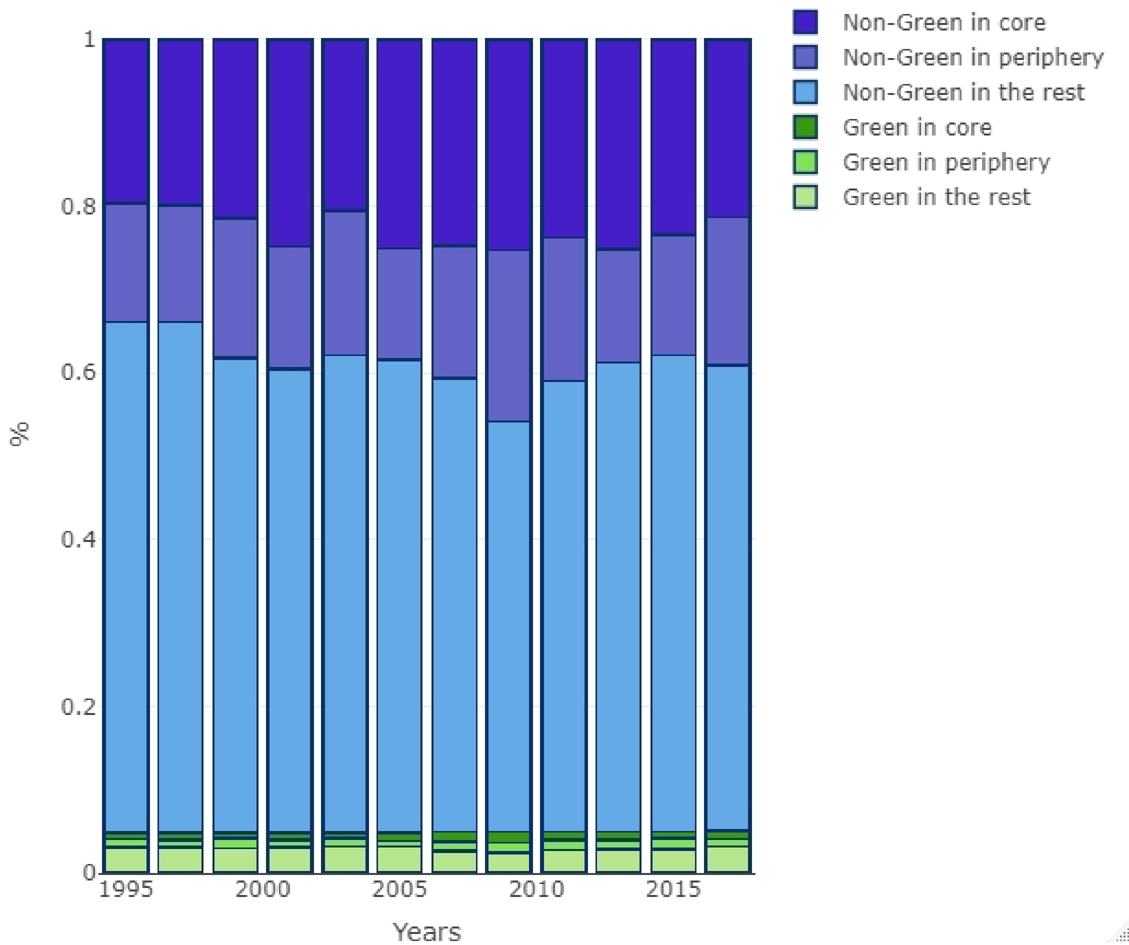
Figure 8: Perron-Frobenius eigenvalue




*Note: The Perron-Frobenius eigenvalue represents the number of feedback loops within the directed product network. A rise in this value signifies an increase in the number of cycles in the network.*

Considering the network's complexity, it is interesting to observe how the size and composition of the core evolve over time. Figure 9 portrays the changing composition of products within the core and periphery of the autocatalytic structure.

Figure 9: Relative share of nodes in the core, periphery and the rest




We observe fluctuations over time across each product class; however, the total number of nodes remains relatively stable.

Table 2 analyzes the evolution of ACS membership over time using the Pearson correlation[12]. All the correlation values are greater than zero.

[12]The Pearson correlation is a statistical measure that gauges the strength and direction of the linear relationship between two variables, ranging from -1 to 1. A value of -1 signifies a perfect negative correlation, 0 denotes no correlation, and 1 signifies a perfect positive correlation.

Table 2: Pearson correlation of product's ACS and core memberships between periods from 1995 to 2019.

| First period | Second period | ACS | core |
|---|---|---|---|
| 1995-1997 | 1997-1999 | 0.32 | 0.58 |
| 1997-1999 | 1999-2001 | 0.33 | 0.59 |
| 1999-2001 | 2001-2003 | 0.32 | 0.59 |
| 2001-2003 | 2003-2005 | 0.36 | 0.58 |
| 2003-2005 | 2005-2007 | 0.35 | 0.60 |
| 2005-2007 | 2007-2009 | 0.39 | 0.62 |
| 2007-2009 | 2009-2011 | 0.34 | 0.58 |
| 2009-2011 | 2011-2013 | 0.34 | 0.59 |
| 2011-2013 | 2013-2015 | 0.38 | 0.62 |
| 2013-2015 | 2015-2017 | 0.42 | 0.64 |
| 2015-2017 | 2017-2019 | 0.38 | 0.61 |
| 2017-2019 | 1995-1997 | 0.29 | 0.44 |

*Note:The membership of a product in an ACS is represented by a binary list, with a value of 1 indicating inclusion and 0 indicating exclusion. These lists are standardized to include only the products that are present in every year (i.e., the common denominator; 4635 goods). Subsequently, a Pearson correlation is computed for all pairs of products in the list. This process is performed for both the core and the ACS.*

This indicates a positive relationship from one ACS membership to the next. More specifically, ACS membership in one period is positively correlated with ACS membership in the preceding period, yielding an average correlation of around 0.35. When we focus specifically on core membership, we observe an even stronger correlation over time, around 0.59, indicating a stronger level of stability in the core relative to the overall ACS.

Table 3: Contingency table for ACS membership

| | | 2017-2019 | | | | |
|---|---|---|---|---|---|---|
| | | Non-ACS | | ACS | | Total |
| 1995-1997 | | Observed | frequency | Observed | frequency | |
| | Non-ACS | 2078 | 0.70 | 889 | 0.30 | 2967 |
| | ACS | 671 | 0.40 | 997 | 0.60 | 1668 |
| | Total | 2749 | 0.59 | 1886 | 0.41 | 4635 |

*Note: observed corresponds to the number of products from each of the categories in 1995-1997 and where they are in 2017-2019. Frequency is the share of products, from each category in 1995-1997, that are in the ACS and non-ACS categories of 2017-2019.*

Products that are part of the ACS during the initial period of 1995-1997 have a probability of approximately 60% of remaining members of the ACS during the final period of 2017-2019. Conversely, products that are not part of the ACS in the first period have a higher probability of around 70% of remaining outside the ACS in the last period. Consequently, the likelihood of ACS members exiting from the set is approximately 40%, whereas non-ACS products have a probability of about 30% of entering the set during the specified time frame.

Table 4: Contingency table for core membership

| | | 2017-2019 | | | | |
|---|---|---|---|---|---|---|
| | | Non-core | | Core | | Total |
| 1995-1997 | | Observed | Frequency | Observed | Frequency | |
| | Non-Core | 3205 | 0.87 | 467 | 0.13 | 3672 |
| | Core | 402 | 0.42 | 561 | 0.58 | 963 |
| | Total | 3607 | 0.78 | 1028 | 0.22 | 4635 |

*Note: observed corresponds to the number of products from each of the categories in 1995-1997 and where they are in 2017-2019. Frequency is the share of products, from each category in 1995-1997, that is in the core and non-core categories in 2017-2019.*

These results are reinforced for core nodes. Products that are part of the core during the initial period have a probability of approximately 58% of persisting as members of the core during the final period. Conversely, products that are not part of the core in the first period have a higher probability of around 87% of remaining outside the core in the last period. Therefore, the likelihood of core members exiting from the set is approximately 42%, whereas non-core products have a probability of about 13% of entering the set during the specified time frame.

This remarkable stability in the autocatalytic structure has significant environmental policy implications. In particular, it highlights that introducing more green goods into the ACS may not be a feasible strategy, given the robust stability of the network.

In Table 5, we show the total number and the distribution of green and non-green products for the period 2017-2019. The non-green goods tend to be in the core, while green goods are more likely to be in the periphery.

Table 5: Average distribution of products across the core, periphery and non-ACS 1995-2019

| | Mean number of products | | | Share of the product typology | | |
|---|---|---|---|---|---|---|
| | Non-green | Green | Total | Non-green | Green | Total |
| Core | 1127.83 | 44.67 | 1172.50 | 24% | 18% | 24% |
| Periphery | 779.25 | 51.00 | 830.25 | 17% | 21% | 17% |
| Non ACS | 2774.92 | 145.92 | 920.83 | 59% | 60% | 59% |
| Total | 4682 | 241.58 | 4923.58 | 100% | 100% | 100% |

As shown in Table 2, we find a strong correlation between nodes belonging to the core over time, indicating that more than half of the products in the core are relevant drivers of catalyzation over the period analyzed. This reinforces the idea that green goods are less likely to be drivers of the diversification process. Therefore, a combination of green and non-green goods is necessary to optimize the process of diversification while also striving towards environmental sustainability.

#### 4.3.2 Measuring the effect of the ACS on export volume

As established previously, a country can rely on non-green goods to develop the required capabilities to produce specific green goods. This final subsection aims to establish whether belonging to an ACS, and therefore benefiting from its multiple feedback loops, can increase a country's ability to export a given good.

ACS is defined differently in this subsection. In the previous sections, the ACS detection algorithm was applied to the global network. In this subsection, the ACS detection algorithm is applied to a subset of the network, namely the domestic product space.

The domestic product space is a subset of the product space. A country's domestic product space is based on the following elements: Nodes are products that are exported by the country with an RCA greater than 1. A link exists when the products at the origin and destination of the catalyzing relationship are nodes (i.e., both the origin and destination countries export the product with RCA>1). The network of nodes and links represents the multiple feedback loops that reinforce goods already produced by a given country.

To gain a deeper insight into the implications of the directed product space on development, we have mapped the positions of each of the BRICS countries (Brazil, Russia, India, China, South Africa) within the directed product space. The result of this exercise is presented in the Appendix A.5. While maintaining the network's structure, we have highlighted the products that these countries are currently exporting.

The hypothesis we aim to validate in this section is whether the ACS detected in the domestic product space reinforces a country's level of export of a given product. If there is a real effect, it means that having clusters of related industries in a country is important for the export of a good. The economic implications of this would be that policies aimed at maintaining a cluster of industries related to a given green good would be preferable to simply acquiring specific capabilities. If the ACS structure has no impact, then the ACS members should not exhibit significantly higher values of RCA in the network than the non-members. Similarly, goods in the core and the periphery should not exhibit significantly higher values either.

We therefore apply the ACS detection algorithm to the domestic product space. It provides us with multiple indicators that are country- and product-specific. ACSlocal measures whether a product is part of the local ACS of a country. Corelocal and Perilocal indicate whether the product is part of the core or the periphery of the ACS of that given country.

We then use a log-log fixed-effect model to measure the effect.

Our dependent variable is a country's continuous export level of a given good at the 6-digit level. To test the importance of belonging to an ACS, we estimated the following equations:

$$\begin{aligned} log(Exp_{c,p,t}) =& \alpha_1 ACSlocal_{c,p,t} + \alpha_2 Maxprox_{c,p,t} \\ &+ \alpha_3 log(GDP_{c,p,t}) + \epsilon_t + C_c + Y_t \end{aligned} \tag{6}$$

$$\begin{aligned} log(Exp_{c,p,t}) =& \alpha_1 Corelocal_{c,p,t} + \alpha_2 Perilocal_{c,p,t} \\ &+ \alpha_2 Maxprox_{c,p,t} + \alpha_3 log(GDP_{c,p,t}) + \epsilon_t + C_c + Y_t \end{aligned} \tag{7}$$

Equation 6 expresses the volume of export of a good as a function of its position as a local ACS member within the directed product space of a country ($ACSlocal_{c,p,t}$), the similarity that the product has with other goods in the country's export basket ($Maxprox_{c,p,t}$), and the size of the country's economy ($log(GDP_{c,p,t})$). $c$ represents the index for countries, $p$ denotes the index for products, and $t$ is the index for time. Equation 7 is the same equation but separates $ACSlocal_{c,p,t}$ into two other variables, $Corelocal_{c,p,t}$ and $Perilocal_{c,p,t}$. C is a country fixed effect and Y is a time fixed effect.

$$\begin{aligned} log(Exp_{c,p,t}) =& \alpha_1 ACSlocal_{c,p,t} + \alpha_2 Maxprox_{c,p,t} \\ &+ \alpha_3 log(GDP_{c,p,t}) + \alpha_4 log(Neighborprox_{c,p,t}) \\ &+ \alpha_5 Sophistication_{p,t} + \alpha_6 log(pop_{c,t}) \\ &+ \alpha_7 log(RCA_{c,p,t}) + \alpha_8 Green_p + \epsilon_t + C_c + Y_t \end{aligned} \tag{8}$$

$$\begin{aligned} log(Exp_{c,p,t}) =& \alpha_1 Corelocal_{c,p,t} + \alpha_2 Perilocal_{c,p,t} \\ &+ \alpha_2 Maxprox_{c,p,t} + \alpha_3 log(GDP_{c,t}) \\ &+ \alpha_4 log(Neighborprox_{c,p,t}) + \alpha_5 Sophistication_{p,t} \\ &+ \alpha_6 log(pop_{c,t}) + \alpha_7 log(RCA_{c,p,t}) + C_c + Y_t \\ &+ \alpha_8 Green_p + \epsilon_t \end{aligned} \tag{9}$$

Equation 8 and Equation 9 take the previous equations and add additional variables. $Maxprox_{c,p,t}$ and $Neighborprox_{c,p,t}$ are measures of a country's opportunity to export a good based on other goods exported by the country and its neighbors. $Green_p$ and $Sophistication_{p,t}$ provide information on a product's characteristics. $log(FDI_{c,t})$, $log(pop_{c,t})$ and $Tradeopenness_{c,p,t}$ provide information on the country's characteristics. All variables are described in the appendix in Table 8.

Panel regression was applied to all goods and to two subgroups, only the green goods, and only the non-green goods. Table 6 provides the results of these regressions.

Belonging to the ACS has a significantly positive effect on the level of exports. In the shorter form of the equation, the effect of being part of the ACS is stronger than in its long-form version. The long-form equation captures effects that the short-form misses; additional variables measuring the sophistication of the good, the proximity of the good to the export basket of neighboring countries, and the country's current RCA level all explain part of this effect.

Moreover, products in the core exhibit higher export levels than those in the periphery, but only in the short version of the equations. Surprisingly, once the additional variables are added, the effect is reversed, and the periphery exhibits higher levels of exports than the core. Both, however, have a significant positive effect on export levels.

Finally, $GDP_{c,t}$ has a positive effect on exports. A country also has a higher chance of exporting a given product if the country's neighbors are already producers of that good. The RCA level of the country also has a positive effect on the level of exports of the good. The sophistication and population size have a negative effect on the export level of the good. It is worth noting that green goods display higher exports than non-green goods.

Table 6: Regression Results

| | Dependent variable: log(exp) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All goods | | | | Green goods | | | | Non-Green goods | | | |
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| ACSlocal | 2.509*** | | 0.248*** | | 2.082*** | | 0.179*** | | 2.552*** | | 0.253*** | |
| | (0.005) | | (0.005) | | (0.025) | | (0.022) | | (0.006) | | (0.005) | |
| Corelocal | | 2.644*** | | 0.200*** | | 2.204*** | | 0.078** | | 2.701*** | | 0.211*** |
| | | (0.007) | | (0.006) | | (0.043) | | (0.036) | | (0.007) | | (0.006) |
| Perilocal | | 2.289*** | | 0.318*** | | 2.019*** | | 0.229*** | | 2.299*** | | 0.318*** |
| | | (0.009) | | (0.007) | | (0.031) | | (0.026) | | (0.009) | | (0.007) |
| maxprox | 6.775*** | 6.753*** | 2.223*** | 2.228*** | 10.230*** | 10.238*** | 4.063*** | 4.056*** | 6.569*** | 6.544*** | 2.114*** | 2.119*** |
| | (0.015) | (0.015) | (0.013) | (0.013) | (0.061) | (0.061) | (0.055) | (0.055) | (0.015) | (0.015) | (0.013) | (0.013) |
| log(real_gdp) | 0.747*** | 0.749*** | 0.643*** | 0.642*** | 0.865*** | 0.865*** | 0.754*** | 0.754*** | 0.742*** | 0.744*** | 0.633*** | 0.632*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.003) | (0.003) | (0.004) | (0.004) | (0.001) | (0.001) | (0.001) | (0.001) |
| log(Neighborprox) | | | 0.442*** | 0.442*** | | | 0.232*** | 0.232*** | | | 0.450*** | 0.450*** |
| | | | (0.002) | (0.002) | | | (0.007) | (0.007) | | | (0.002) | (0.002) |
| Sophistication | | | 0.268*** | 0.268*** | | | 0.259*** | 0.259*** | | | 0.268*** | 0.268*** |
| | | | (0.001) | (0.001) | | | (0.003) | (0.003) | | | (0.001) | (0.001) |
| log(pop) | | | -0.099*** | -0.099*** | | | -0.159*** | -0.159*** | | | -0.093*** | -0.093*** |
| | | | (0.001) | (0.001) | | | (0.004) | (0.004) | | | (0.001) | (0.001) |
| log(RCA) | | | 1.697*** | 1.699*** | | | 2.050*** | 2.051*** | | | 1.682*** | 1.684*** |
| | | | (0.002) | (0.002) | | | (0.010) | (0.010) | | | (0.002) | (0.002) |
| Green | | | 0.974*** | 0.972*** | | | | | | | | |
| | | | (0.005) | (0.005) | | | | | | | | |
| Observations | 2,540,199 | 2,540,199 | 2,500,444 | 2,500,444 | 163,065 | 163,065 | 160,001 | 160,001 | 2,377,134 | 2,377,134 | 2,340,443 | 2,340,443 |
| $R^2$ | 0.358 | 0.359 | 0.580 | 0.580 | 0.506 | 0.506 | 0.679 | 0.679 | 0.353 | 0.353 | 0.573 | 0.573 |
| Adjusted $R^2$ | 0.358 | 0.359 | 0.580 | 0.580 | 0.506 | 0.506 | 0.679 | 0.679 | 0.353 | 0.353 | 0.573 | 0.573 |
| F Statistic | 473,010.800*** | 355,194.000*** | 431,764.000*** | 383,838.800*** | 55,569.300*** | 41,683.150*** | 48,248.180*** | 42,221.890*** | 432,337.200*** | 324,778.500*** | 448,488.100*** | 392,468.300*** |
| | (df = 3; 2540185) | (df = 4; 2540184) | (df = 8; 2500425) | (df = 9; 2500424) | (df = 3; 163051) | (df = 4; 163050) | (df = 7; 159983) | (df = 8; 159982) | (df = 3; 2377120) | (df = 4; 2377119) | (df = 7; 2340425) | (df = 8; 2340424) |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Each model includes a high number of observations (over 2.5 million) for the version with all observations included. The green goods subgroup includes around 160,000 observations. Each model demonstrates strong explanatory power in its long form. For all observations, the R2 value corresponds to 0.58. The model has the strongest explanatory power in the green goods subgroup, with an R2 value of 0.679. The F-statistics are also highly significant, indicating the overall validity of each model.

To sum up, having a product within a country's ACS is positively correlated with a higher value of exports. It is therefore important to consider having a cluster of domestic industries instead of focusing on a single green industry. This domestic cluster of industries is important as it enables catalyzing feedback loops. In contrast to the framework presented in Jain and Krishna (1998), the periphery has a stronger positive effect than the core in terms of the size of exports.

Finally, when comparing the two subgroups, the effect of the ACS on the export of green goods is weaker than the effect on non-green goods. This is true both as an aggregate value and when separated between core and periphery. This reinforces the idea that the catalyzation effect is lower for green goods than for non-green goods. The significance of the effect however remains high both for green and non-green goods.

# 5 Concluding Remarks

In conclusion, the global imperative of addressing climate change poses significant challenges, especially for developing countries most adversely impacted by its effects. The domestic manufacture of environmental products is important to reinforce the ability of developing countries to create tailored policies to address climate change. Nevertheless, considerable obstacles persist, mainly in the shape of constrained financial and technological assets, insufficient infrastructure, and lack of a skilled workforce.

This paperhas investigated a specific type of constraint that affects the set of capabilities of a given country. To do this, we have modeled the inter-industry relationships in the form of a directed product space. In this directed product space, we have identified environmental products. Furthermore, we have described the position of these green goods in the network, comparing their direct and indirect catalyzation power with their non-green good counterparts. We have also analyzed the cycles in the network, especially those that include both green and non-green goods. Finally, we have accounted for the multiple overlapping cycles in the network by modeling the ACS.

We showed that green goods benefit from fewer self-reinforcing effects than their non-green counterparts. This effect translates into a lower effect of the ACS on export level. This implies that green industrial policies are necessary to facilitate the emergence of green products. Indeed, our results suggest that left to their own devices, non-green products will be preferred by market forces. We also showed that green goods provide fewer opportunities for diversification compared to other products. Governments in developing countries that prioritize economic development may be less interested in the adoption of green products. Therefore, several incentives should be created by the international community to push for the development of green goods in developing countries.

We also described the interaction between green and non-green goods. Green goods are more likely to be in a feedback loop with non-green goods than with other green goods. When this occurs, the non-green good is more likely to emerge first and initiate this positive feedback loop.

There are three main avenues of research following this paper. First, additional research needs to be done to understand why peripheral nodes have a stronger catalytic effect on exports than those in the core, as they are supposed to receive fewer feedback loops than in the core. Second, this paper studies the relationship between green and non-green goods. Measuring the relationship between goods that have a negative effect on the environment and green goods could improve policy-makers' choices between the two types of industries. Lastly, more research is needed to investigate whether economic development takes precedence over the development of green goods. Such research can be conducted by simulating the evolution of a country's export basket within the directed product space.

The challenge lies in striking a balance between fostering green product development and encouraging economic diversification. To address this, we propose a multi-faceted approach. The adoption of green products should be integrated into a broader industrialization strategy, rather than focusing solely on the optimal path toward the adoption of green goods. A more strategic approach might be to identify a cluster of industries for specialization and subsequently encourage the adoption of related green industries. More precisely, developing countries should aim to diversify their industries as a means to acquire a large set of capabilities.

In countries that have the required capabilities for the development of green industries, policies should be put in place to enhance incentives for both policy-makers and businesses to promote the growth of green industries. This could include rewarding sustainable practices, providing subsidies for green product development, or implementing mechanisms to provide monetary and non-monetary support to foster green industry development. These mechanisms could operate at various levels: national politicians might drive domestic incentives, while international players like institutions or other countries could provide external investment. This holistic approach could foster sustainable development without compromising the broader goal of industrial diversification.

Our research also underscores the complexity of the transition to sustainability, which cannot be achieved by merely substituting non-green goods with green ones. With the insights gained from our analysis, we hope to contribute to the ongoing discourse on managing the delicate balance between economic development and environmental care, and inspire further research to address the multifaceted challenges of sustainability.

# References


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

Alhaddad, H., Talebzadehhosseini, S., and Garibay, I. (2023). *Green Sector Space: The Evolution and Capabilities Spillover of Economic Green Sectors in the United States*, pages 502–513.

APEC (2012). *Leaders Declaration Annex C.*

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage1. *The Manchester School*, 33(2):99–123.

Balineau, G. and De Melo, J. (2013). Removing barriers to trade on environmental goods: an appraisal. *World Trade Review*, 12(4):693–718.

Barbieri, N., Consoli, D., Napolitano, L., Perruchas, F., Pugliese, E., and Sbardella, A. (2022). Regional technological capabilities and green opportunities in europe. *The Journal of Technology Transfer*, 48:1–30.

Barbieri, N., Marzucchi, A., and Rizzo, U. (2023). Green technologies, interdependencies, and policy. *Journal of Environmental Economics and Management*, page 102791.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.

Capasso, M., Hansen, T., Heiberg, J., Klitkou, A., and Steen, M. (2019). Green growth – a synthesis of scientific findings. *Technological Forecasting and Social Change*, 146:390–402.

CEPII (2023). `http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=17`. Accessed: 2023-03-28.

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*, 1(3):140216.

Cunzo, F., Petri, A., Zaccaria, A., and Sbardella, A. (2022). The trickle down from environmental innovation to productive complexity. *Scientific Reports*, 12:22141.

de Melo, J. and Solleder, J.-M. (2017). Non-tariff measures and other trade measures for the environment in asean. Technical report, FERDI Working Paper.

Environmental Business International inc. (2023). https://ebionline.org/. Accessed: 2023-03-28.

Fankhauser, S., Bowen, A., Calel, R., Dechezleprêtre, A., Grover, D., Rydge, J., and Sato, M. (2013). Who will win the green race? in search of environmental competitiveness and innovation. *Global Environmental Change*, 23(5):902–913.

Feenstra, R. C., Lipsey, R. E., Deng, H., Ma, A. C., and Mo, H. (2005). World trade flows: 1962-2000. Working Paper 11040, National Bureau of Economic Research.

Ferraz, D., Sartori Falguera, F., Mariano, E., and Hartmann, D. (2021). Linking economic complexity, diversification, and industrial policy with sustainable development: A structured literature review. *Sustainability*.

Fischer, C. and Newell, R. G. (2008). Environmental and technology policies for climate mitigation. *Journal of Environmental Economics and Management*, 55(2):142–162.

Fraccascia, L., Giannoccaro, I., and Albino, V. (2018). Green product development: What does the country product space imply? *Journal of cleaner production*, 170:1076–1088.

Gaulier, G., Martin, J., Méjean, I., and Zignago, S. (2008). International trade price indices. Working Papers 2008-10, CEPII.

Hamwey, R., Pacini, H., and Assunção, L. (2013). Mapping green product spaces of nations. *The Journal of Environment & Development*, 22(2):155–168.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., and Yıldırım, M. A. (2013). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. The MIT Press.

Hausmann, R. and Klinger, B. (2006). Structural transformation and patterns of comparative advantage in the product space.

Hickel, J. (2018). Is it possible to achieve a good life for all within planetary boundaries? *Third World Quarterly*, 40(1):18–35.

Hickel, J. and Kallis, G. (2020). Is green growth possible? *New Political Economy*, 25(4):469–486.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837):482–487.

Huberty, M. and Zachmann, G. (2011). Green exports and the global product space-Prospects for EU industrial policy. Working Papers 556, Bruegel.

Jackson, T. (2009). *Prosperity Without Growth? The Transition to a Sustainable Economy.*

Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014). Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLOS ONE*, 9(6):1–12.

Jaffe, A., Newell, R., and Stavins, R. (2005). A tale of two market failures: Technology and environmental policy. *Ecological Economics*, 54(2-3):164–174.

Jain, S. and Krishna, S. (1998). Autocatalytic sets and the growth of complexity in an evolutionary model. *Phys. Rev. Lett.*, 81:5684–5687.

Jain, S. and Krishna, S. (2003). Graph theory and the evolution of autocatalytic networks. *Handbook of Graphs and Networks*, page 355.

Lapatinas, A., Garas, A., Boleti, E., and Kyriakou, A. (2019). Economic complexity and environmental performance: Evidence from a world sample. MPRA Paper 92833, University Library of Munich, Germany.

Mealy, P. and Teytelboym, A. (2020). Economic complexity and the green economy. *Research Policy*, page 103948.

Müller, V. and Eichhammer, W. (2023). Economic complexity of green hydrogen production technologies - a trade data-based analysis of country-specific industrial preconditions. *Renewable and Sustainable Energy Reviews*, 182.

Napolitano, L., Evangelou, E., Pugliese, E., Zeppini, P., and Room, G. (2018). Technology networks: the autocatalytic origins of innovation. *Royal Society Open Science*, 5(6):172445.

Neffke, F. and Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3):297–316.

OCDE (2011). *Towards Green Growth.*

Ortiz-Ospina, E. and Beltekian, D. (2018). How and why should we study 'economic complexity'? https://ourworldindata.org/how-and-why-econ-complexity.

Popp, D., Newell, R., and Jaffe, A. (2009). Energy, the environment, and technological change. NBER Working Papers 14832, National Bureau of Economic Research, Inc.

Pugliese, E., Cimini, G., Patelli, A., Zaccaria, A., Pietronero, L., and Gabrielli, A. (2019). Unfolding the innovation system for the development of countries: coevolution of science, technology and production. *Scientific reports*, 9(1):1–12.

Quan, G., Tie-Ju, M., and Chun-Jie, C. (2017). Evolution of world exports: Based on mutualistic environment within a product space. In *2017 International Conference on Management Science and Engineering (ICMSE)*, pages 388–392.

Rodrik, D. (2014). Green industrial policy. *Oxford Review of Economic Policy*, 30(3):469–491.

Romero, J. P. and Gramkow, C. (2021). Economic complexity and greenhouse gas emissions. *World Development*, 139:105317.

Saracco, F., Di Clemente, R., Gabrielli, A., and Squartini, T. (2015). Randomizing bipartite networks: The case of the world trade web. *Scientific Reports*, 5.

Sauvage, J. (2014). The stringency of environmental regulations and trade in environmental goods. Working paper, Paris : OECD Publishing.

Savona, M. and Ciarli, T. (2019). Structural Changes and Sustainability. A Selected Review of the Empirical Evidence. *Ecological Economics*, 159(C):244–260.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. (2013). Economic complexity: Conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control*, 37(8):1683–1691.

Talebzadehhosseini, S. and Garibay, I. (2021). The interaction effects of technological innovation and path-dependent economic growth on countries overall green growth performance. *Journal of Cleaner Production*, 333:130134.

Talebzadehhosseini, S., Scheinert, S., and Garibay, I. (2019). Global transitioning towards a green economy: Analyzing the evolution of the green product space of the two largest world economies. pages 633–644.

Teece, D. J., Rumelt, R., Dosi, G., and Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior Organization*, 23(1):1–30.

UNFCCC (2015). *Paris Agreement. 21st Conference of the Parties*. Paris: UNFCCC.

United Nations (2012). The world we want.

World Bank (2012). Inclusive green growth: The pathway to sustainable development.

WTO (2011). Report by the chairman to the trade negotiations committee, committee on trade and environment in special session, doc. tn/te/20. Technical report.

Zaccaria, A., Cristelli, M., Tacchella, A., and Pietronero, L. (2014). How the taxonomy of products drives the economic development of countries. *PloS one*, 9(12):e113770.

# A Appendix

## A.1 Additional Tables

Table 7: Number of green and non-green goods in each HS92 2-digit category

| Sectors | Non-green | Green |
|---|---|---|
| Animal & Animal Products | 194 | 0 |
| Chemicals & Allied Industries | 759 | 1 |
| Foodstuffs | 181 | 0 |
| Footwear / Headgear | 55 | 0 |
| Machinery / Electrical | 659 | 105 |
| Metals | 563 | 25 |
| Mineral Products | 151 | 0 |
| Miscellaneous | 338 | 49 |
| Plastics / Rubbers | 182 | 7 |
| Raw Hides, Skins, Leather, & Furs | 74 | 0 |
| Stone / Glass | 176 | 12 |
| Textiles | 800 | 9 |
| Transportation | 95 | 37 |
| Vegetable Products | 323 | 0 |
| Wood & Wood Products | 225 | 3 |

Table 8: Variable list

| Variable name | Type | Definition | Source |
|---|---|---|---|
| $Export_{c,p,t}$ | Continuous | Value of export of the product (p) by country (c) at time (t) | BACI database |
| $ACSlocal_{c,p,t}$ | Binary | Dummy variable indicating whether a product (p) is part of the local ACS of the country (c) at time (t) | Calculation of the authors |
| $Corelocal_{c,p,t}$ | Binary | Dummy variable showing if a product is in the core of the local ACS of the country (c) at time (t) | Calculation of the authors |
| $Perilocal_{c,p,t}$ | Binary | Dummy variable showing if a product is in the periphery of the local ACS of the country (c) at time (t) | Calculation of the authors |
| $Maxprox_{c,p,t}$ | Continuous | Proximity between a given product (p) and the nearest product exported by the same country (c) at time (t) | Calculation of the authors |
| $Green_p$ | Binary | Dummy variable equal to 1 if the product is "green"and 0 otherwise. | CLEG |
| $Neighborprox_{c,p,t}$ | Continuous | Average proximity at time (t) between a given product (p) and all the other goods exported by the geographical neighbors of the country (c) | Calculation of the authors |
| $Sophistication_{p,t}$ | Continuous | Measure of a product sophistication based on Fitness-complexity method by Tacchella et al. (2013) | Calculation of the authors |
| $Real_gdp_{c,t}$ | Continuous | GDP PPP of the country (c) at time (t) | CHELEM database |
| $pop_{c,p,t}$ | Continuous | Population of the country (c) at time (t). | CHELEM database |
| $RCA_{c,p,t}$ | Continuous | RCA of the country (c) in the production of the product (p) at the time (t) | Calculation of the authors |

*Note: The proximity matrices are based on the proximity network introduced by Hausmann and Klinger (2006) see Appendix A.2.1*

## A.2 Formula

### A.2.1 Hausmann's Undirected Product Space

The Product Space, as introduced by Hausmann and Klinger (2006), is a network representation of the set of products exported by different countries. Each product is a node in the network, and the edges between nodes represent the probability of two goods being exported by the same country, or in other words, the conditional probability that, given a country exports one product, it also exports the other.

The proximity measure between two products $p$ and $p'$, based on the $M_{c,p}$, is then computed as the minimum of the conditional probabilities that, given a country has comparative advantage in product $p$, it also has a comparative advantage in product $p'$, and vice versa.

$$\phi_{pp'} = \min \left\{ \frac{\sum_c M_{cp} M_{cp'}}{\sum_c M_{cp}}, \frac{\sum_c M_{cp} M_{cp'}}{\sum_c M_{cp'}} \right\} \tag{10}$$

The higher is the value of $\phi_{pp'}$ the higher is the probability of $p$ and $p$' to be co-exported. Based on the probability of co-exportation of a good, the proximity index serves as a proxy for the similarity in terms of capability requirements between two products.

### A.2.2 Sophistication Index

The Sophistication Index originating from Tacchella et al. (2013), serves as a measure of sophistication for both products and countries. The product sophistication serves as a proxy for the diversity of capabilities required for producing a given good. This complexity requires advanced technologies, specialized skills, infrastructure, or an advanced division of labor.

Similarly, country sophistication measures a country's ability to export sophisticated products. This ability is correlated with the country's technological advancement and division of labor, as it demonstrates that the country has a large set of capabilities that it can combine to produce complex goods.

The Sophistication Index, therefore, is a valuable tool for measuring a country's comparative advantage and potential for economic growth. For a detailed explanation of the Sophistication Index, see (Tacchella et al., 2013).

$$\begin{cases} \tilde{F}_c^n = \sum_p M_{cp} Q_{cp}^{(n-1)} \\ \tilde{Q}_p^n = \frac{1}{\sum_c M_{cp} \frac{1}{F_{cp}^{(n-1)}}} \end{cases} \tag{11}$$

$$\begin{cases} F_c^n = \frac{\tilde{F}_c^n}{\langle \tilde{F}_c^n \rangle} \\ Q_c^n = \frac{\tilde{Q}_p^n}{\langle \tilde{Q}_p^n \rangle} \end{cases} \tag{12}$$

The Fitness-Complexity algorithm proposed by Tacchella et al. is an iterative method that quantifies the sophistication of countries and the complexity of products.

Here's a breakdown of the symbols in the equation:

- $F_c^n$: The sophistication of country $c$ at the $n$-th iteration of the algorithm.
- $Q_p^n$: The complexity of product $p$ at the $n$-th iteration of the algorithm.
- $M_{cp}$: A binary matrix where $M_{cp} = 1$ if country $c$ has a revealed comparative advantage (RCA) in product $p$, and $M_{cp} = 0$ otherwise.
- $\tilde{F}_c^n$ and $\tilde{Q}_p^n$: The raw (unnormalized) sophistication of country $c$ and complexity of product $p$ at the $n$-th iteration.
- $\langle \tilde{F}_c^n \rangle$ and $\langle \tilde{Q}_p^n \rangle$: The average raw sophistication and complexity across all countries and products, respectively, at the $n$-th iteration.

The algorithm operates as follows:

1. In the first set of equations, $\tilde{F}_c^n$ is computed as the sum of the complexities of the products that country $c$ has an RCA in, and $\tilde{Q}_p^n$ is the reciprocal of the sum of reciprocals of the sophistications of the countries that have an RCA in product $p$.

2. In the second set of equations, the raw sophistication and complexity values are normalized by dividing each by their respective averages, to get $F_c^n$ and $Q_p^n$.

3. This process is iterated until the values of $F_c^n$ and $Q_p^n$ diverge.

The intuition behind this algorithm is to calculate the sophistication of a country by looking at the complexities of the products it exports, and to estimate the complexity of a product by considering the sophistication of the countries that export it. A high sophistication score indicates that a country can produce a diverse set of complex products, while a high complexity score suggests that a product is made by a small number of highly fit countries.

### A.2.3 PRODY index

The PRODY index is designed as a proxy to estimate the level of economic development, or more specifically, the GDP per capita, at which a particular good emerges within a country's export basket (**?**). It achieves this by calculating the weighted average GDP per capita of all countries that export the said product. The weights are determined based on each country's Revealed Comparative Advantage (RCA), a concept introduced by Balassa (1965), which is a measure of a country's relative advantage or disadvantage in exporting certain goods.

In essence, the PRODY indicator provides insights into the economic sophistication of a product in the global market. A high PRODY value suggests that the product is primarily exported by countries with high GDP per capita, implying that the production of such goods might require advanced technologies, high-level skills, or substantial infrastructure that are typically found in more developed economies. Conversely, a lower PRODY value may be associated with products exported by less developed countries, potentially reflecting goods that require less complex capabilities for their production.

$$PRODY_k = \sum_j \frac{(x_{jk}/X_j)}{\sum_j (x_{jk}/X_j)} Y_j \tag{13}$$

It computes the expected income level or GDP per capita associated with a particular product, $k$. $j$ denotes a specific country in the set of all countries. $x_{jk}$ is the export value of product $k$ by country $j$. $X_j$ is the total export value of country $j$. $Y_j$ is the GDP per capita of country $j$. The weight assigned to each country's GDP per capita ($Y_j$) in calculating the PRODY value is the Revealed Comparative Advantage (RCA) of country $j$ in exporting product $k$.

The intuition behind this formula is to give more weight to the GDP per capita of countries that have a higher RCA in the particular product. The PRODY indicator, therefore, reflects the income level of the countries that are relatively more successful or competitive in exporting product $k$.

## A.3 Network analysis

Table 9: Network metrics of the Directed Product Space

| Years | 1995-1997 | 2001-2003 | 2011-2013 | 2017-2019 |
|---|---|---|---|---|
| Edge | 25094 | 28615 | 26244 | 22210 |
| Node | 2821 | 3054 | 2933 | 2641 |
| Product | 5015 | 5003 | 4878 | 4647 |
| Average out-degree | 5.004 | 5.72 | 5.38 | 4.779 |
| Average shortest path | 3.922 | 4 | 4.122 | 4.123 |
| Maximum degree | 246 | 250 | 275 | 268 |
| Maximum out-degree | 131 | 133 | 138 | 142 |
| Maximum in-degree | 118 | 123 | 137 | 126 |
| Diameter | 12 | 11 | 12 | 11 |
| Completeness | 0.001 | 0.001 | 0.001 | 0.001 |
| Reciprocity | 0.146 | 0.171 | 0.192 | 0.186 |
| Average clustering coeff | 0.175 | 0.185 | 0.192 | 0.181 |
| Assortativity in degree | 0.256 | 0.231 | 0.258 | 0.26 |

In this section of the appendix, we provide a network analysis of the general structure of the Product Space. The general structure of the network yields important information on the dynamics of the network and the prospects for diversification for a country.

Table 9 presents the general characteristics of the Product Spaces for a few specific periods that are representative of the period studied: (1995-1997), (2001-2003), (2011-2013), and (2017-2019). Comparing the number of nodes in the network with the number of products within BACI shows that half of the products are scattered nodes without significant links.

Since the Product Space is a network, its structure can be analyzed through standard network analysis. Our analysis includes the following network metrics: average out-degree, average shortest path, maximum degree, diameter, completeness, reciprocity, average clustering coefficient, and assortativity.

A link represents the number of catalyzing relationships in the network. If the catalyzation is reciprocal, it will be considered as two separate links. The probability of a catalyzing effect being reciprocal is between 14% and 19%. Nodes correspond to the number of products that are connected to at least one other node.

Using a 0.1% significance level, only around 22210 to 28615 links remain. The completeness index, which is the ratio between the number of links and the total possible number of links, is 0.001 and remains the same throughout the period. This level of completeness indicates that the network is sparse.

The number of edges follows two trends: an increase during the first two periods from 25094 in the period 1995-1997 to 28615 in the period 2001-2003, followed by a decrease to 22210 in the period 2017-2019. This suggests that the network initially increased the number of catalytic relationships and then decreased them in the last two periods. This trend aligns with the evolution of the total number of connected products in the network.

The average out-degree represents the mean count of outgoing links from each node, equating to the average number of different products that a given product catalyzes. On average, a product catalyzes 4 to 6 goods. However, the number of catalyzed goods can vary significantly. The maximum out-degree refers to the node with the highest degree of connectivity; in this network, that value is 142 goods.

While many links have been filtered, the average clustering coefficient, which is the probability of two neighbors of a node also being connected, ranges between 10% and 20% depending on the year. This suggests that some nodes have few or no outgoing links, while others are in a dense region of the network. Both the reciprocity rate and the clustering coefficient indicate that multiple feedback loop effects can occur due to the network's clustered structure.

The assortativity in degree, which is the likelihood of a node being connected to another node with a similar number of connections, shows that goods with multiple connections are more likely to connect with other highly connected nodes. Negative assortativity implies that more connected nodes are likely to connect with nodes that have fewer connections.

Both the average shortest path and the network's diameter[13] indicate that relatively few steps of catalyzed products are needed to connect all the connected

[13]The diameter of a network is the longest shortest possible path in the network.

products in the network. Overall, the network appears to be relatively stable over time, with similar measures for most of the analyzed features across the four time periods.

## A.4 Comparing green and non-green Centrality in the Directed Product Space

We continue our analysis by comparing the relative importance of green goods in the network. Table 10 provides a comparison of network centrality statistics for green and non-green nodes. We will interpret these centrality scores and discuss their economic implications for the development of green products as well as their impact on further diversification.

Table 10: T-test comparing the average centrality between green and non-green goods

| Average centrality | Non-green product | Green product | P-value | |
|---|---|---|---|---|
| In-degree | 8.48 | 6.85 | 0.09 | * |
| Out-degree | 8.56 | 5.28 | 5.29e-05 | *** |
| Betweenness | 3365.41 | 2016.23 | 0.004 | *** |
| Closeness | 0.35 | 0.28 | 0.001 | *** |
| Eigenvector (upstream) | 0.03 | 0.008 | 9.21e-17 | *** |
| PageRank (upstream) | 0.0003 | 0.0002 | 3.00e-11 | *** |
| Eigenvector (downstream) | 0.04 | 0.02 | 0.04 | ** |
| PageRank (downstream) | 0.0003 | 0.0002 | 2.25e-06 | *** |

*Note: 128 green products and 2513 non-green products between 2017 and 2019. Only the nodes connected within the network is considered in this analysis. * $p < 0.1$, **$p < 0.5$, ***$p < 0.01$.*

First, the out-degree centrality measures the number of outgoing connections from a node in a network, indicating the number of products catalyzed by a given product. Non-green goods have higher out-degree centrality on average, playing a significant role in the diversification process. Conversely, the lower average out-degree centrality of green goods signifies fewer diversification opportunities.

When considering indirect connections, similar results are observed. A similar pattern is found in Closeness centrality[14]. Closeness centrality measures the product's average distance to all other reachable goods. On average, non-green goods are closer to other goods, indicating better prospects for diversification compared to green goods. However, this can vary considerably from one product to another. Some green goods have higher centrality than some brown goods.

[14]Corresponding to the inverse of the sum of distances from a node to all other reachable nodes

In-degree centrality counts the number of goods that catalyze a specific good. Higher in-degree centrality for non-green goods indicates that they are more likely to be catalyzed by other goods in the network.

Betweenness centrality measures the importance of a node in a network. It quantifies how often a node lies on the shortest path between other nodes. Goods with higher betweenness centrality play a key role in the development process, serving as necessary paths for the adoption of a wide range of products. Non-green products have higher betweenness scores compared to their green counterparts, making them more crucial for the development of new industries.

Eigenvector centrality and PageRank centrality measure the importance of a node based on the centrality of its neighbors. These measures assign relative scores to all nodes in the network, considering not only the number of connections but also the importance of connected nodes. The main difference between the two is that PageRank centrality includes a dampening factor that reduces the influence of distant nodes.

Upstream eigenvector centrality and upstream PageRank centrality gauge the importance of nodes based on the importance of their catalyzing goods. Higher scores indicate that these goods are more likely to be catalyzed due to the network structure. The difference in average upstream centrality measures shows that non-green goods are catalyzed by more important nodes. Reciprocally, downstream eigenvector centrality and downstream PageRank centrality measure the importance of green goods based on the centrality of the goods they catalyze. The difference in average downstream centrality measures shows that non-green goods are more likely to catalyze important goods in the long run.

The network centrality measures illustrate two key aspects of green goods' position in the Directed Product Space. First, green goods offer fewer diversification opportunities on average compared to non-green counterparts. This presents a trade-off for resource-limited countries when selecting industries for investment. Second, green goods are less frequently catalyzed within the Directed Product Space, indicating challenges to their incorporation into existing industry frameworks or supply chains.

## A.5 BRICS countries position in the Directed Product Space

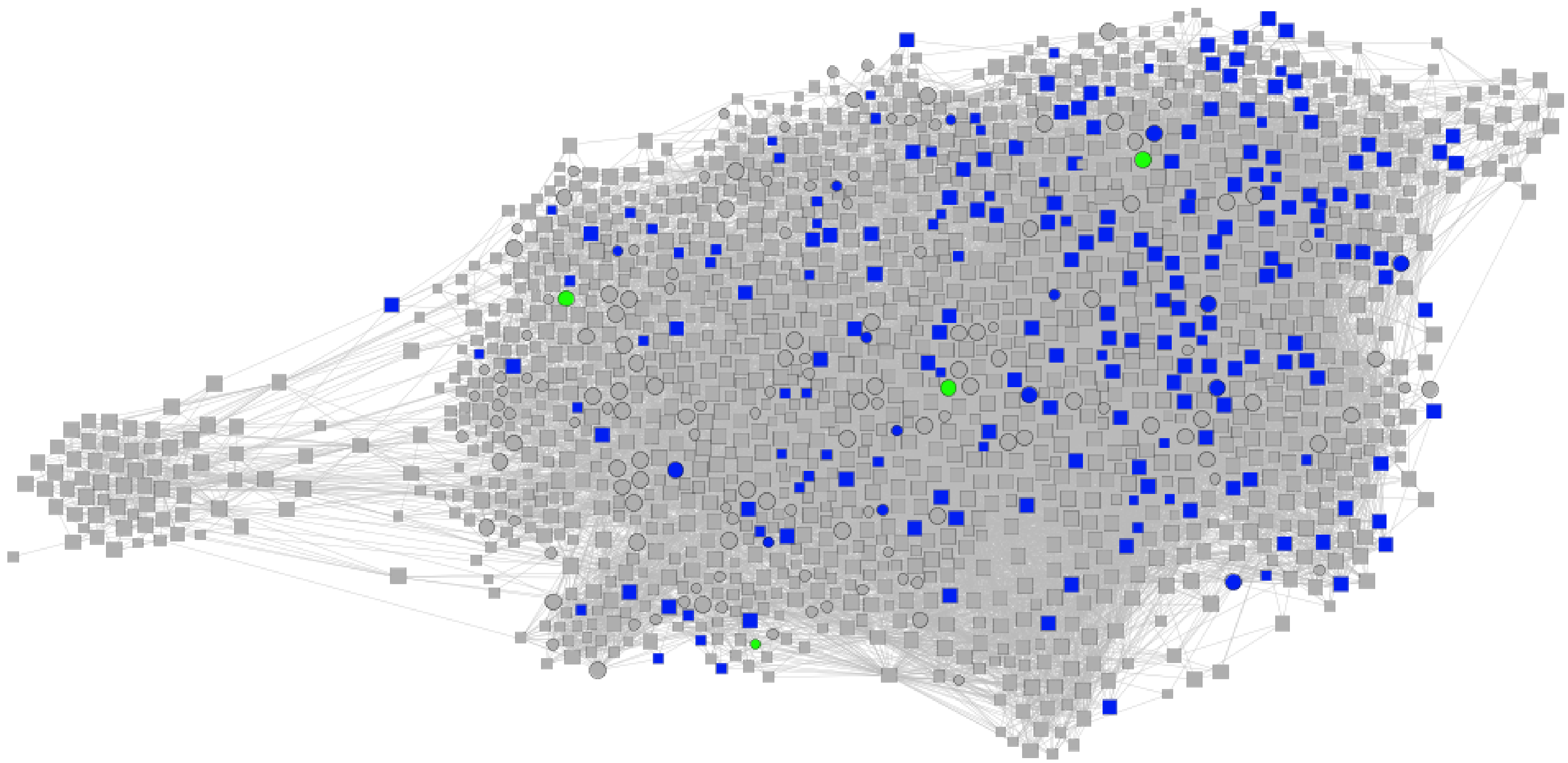
In line with previous findings on the Product Space, each country's position within the Directed Product Space highlights its industrial specialization.

Starting with Russia, two patterns describe its position within the Directed Product Space. The first pattern comprises isolated nodes, and the second consists of nodes concentrated into a cluster. These clustered nodes receive stronger positive feedback loops than the isolated nodes, reinforcing Russia's specialization and its position in the exportation of extractive industries.This reinforcing pattern also increases the likelihood of developing green goods related to its specialization. For the years 2017-2019, only two green products are displayed in the network.

India is a highly diversified country with a diverse array of goods, giving it a strong advantage in catalyzing a diverse range of green goods. China bears similarities to India, having a large variety of exports which affords it the ability to diversify toward many green goods. However, its specialization in normal goods differs, with a tendency toward more complex goods.

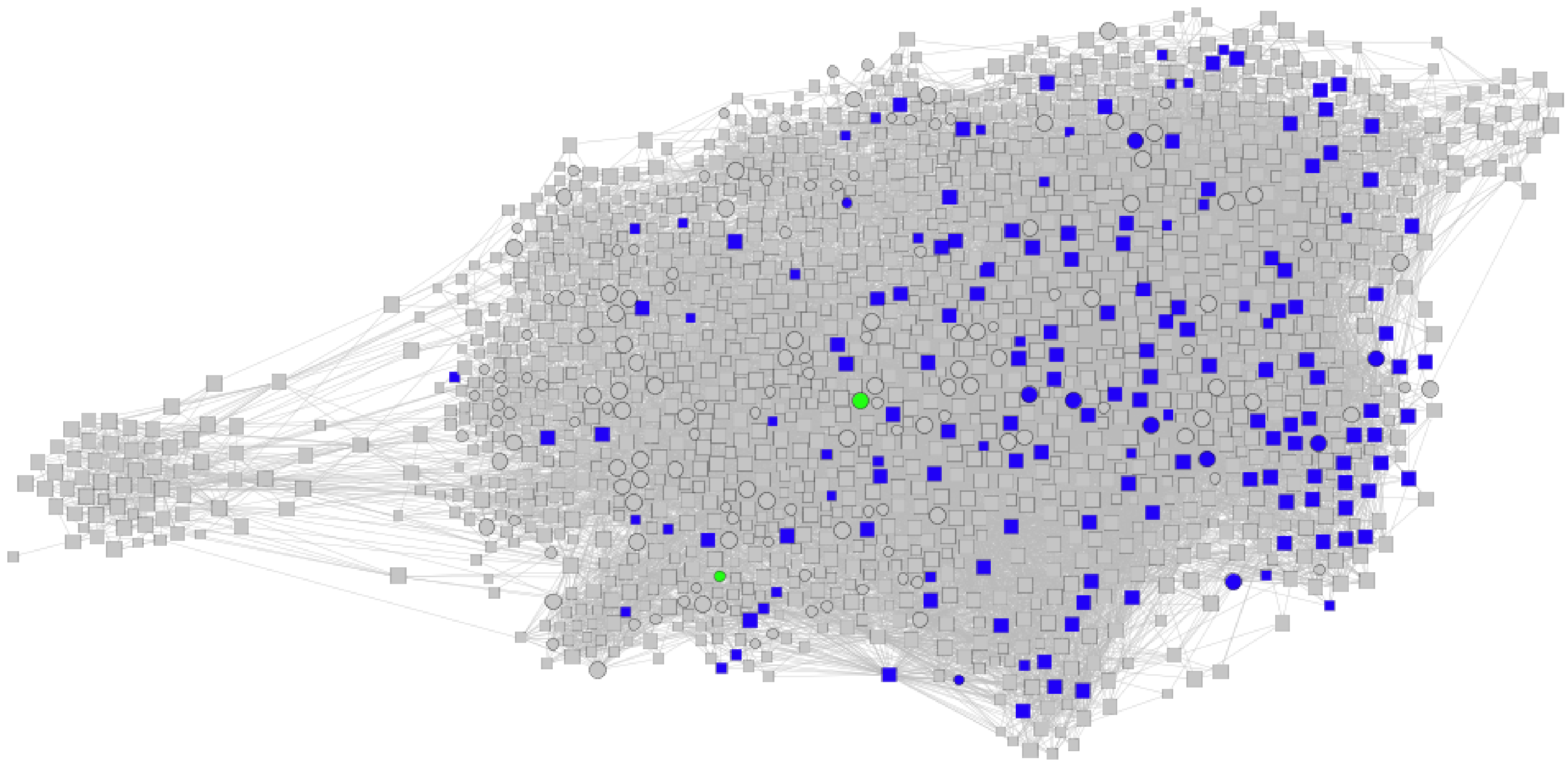
In comparison to other BRICS countries, the goods exported from Brazil are scattered in the network. This scattering can reduce the number of positive feedback loops that reinforce the production of goods, especially green goods. South Africa's position in the ACS resembles that of Russia, featuring both unconnected nodes in the ACS and a group of nodes organized into clusters.

Figure 10: Brazilian Product Space and Autocatalytic Set



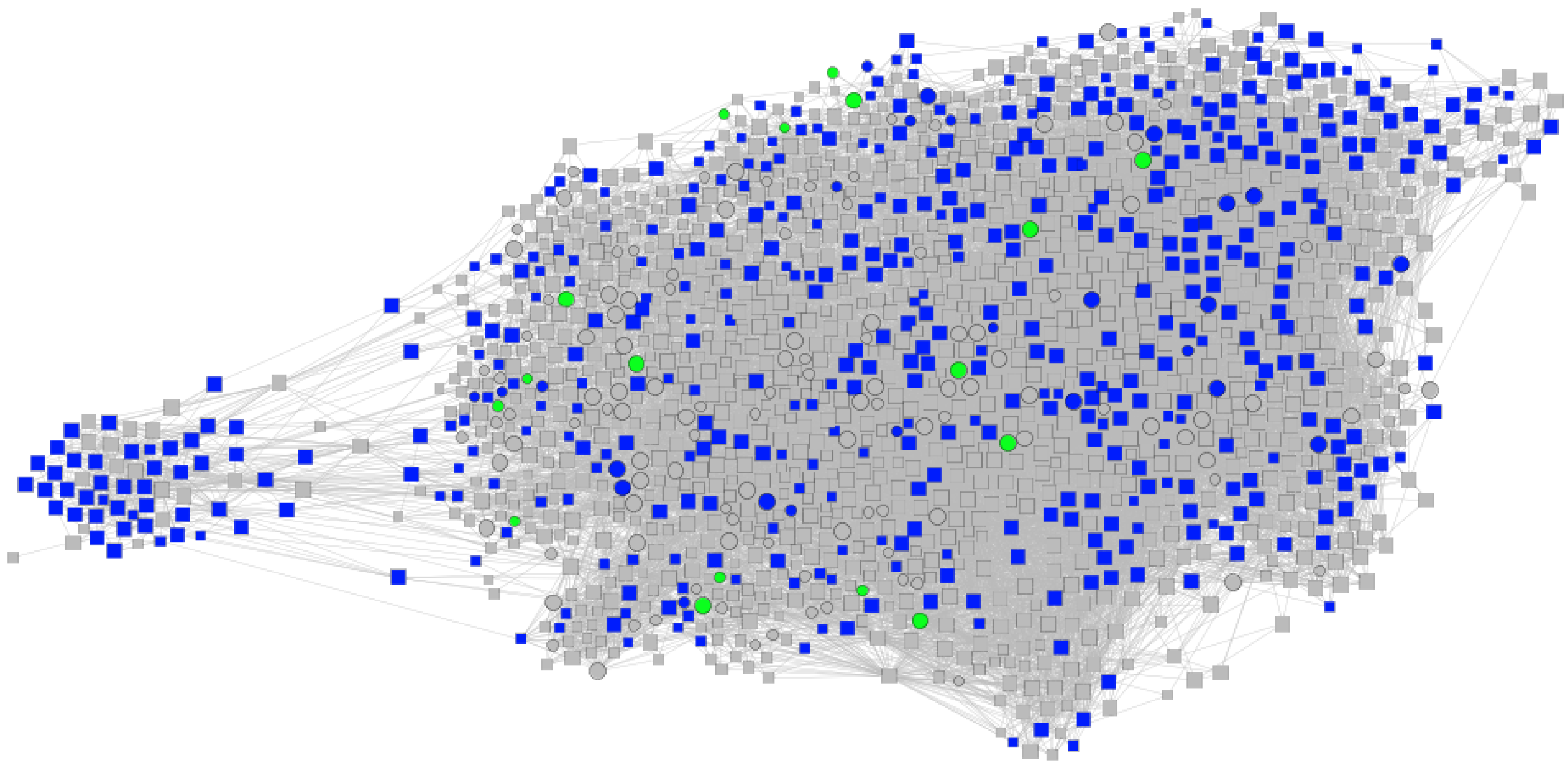
*Note: We model the Brazilian Product Space as a network in which nodes represent products exported by Brazil with an RCA $> 1$, and links signify catalytic relations as represented in the Global Product Space. On this network, we apply the ACS detection algorithm to model the Brazilian Autocatalytic Set. Blue = Non-green products, Green = Green products, Gray = Products exported with an RCA $< 1$.*

Figure 11: Russian Product Space and Autocatalytic Set



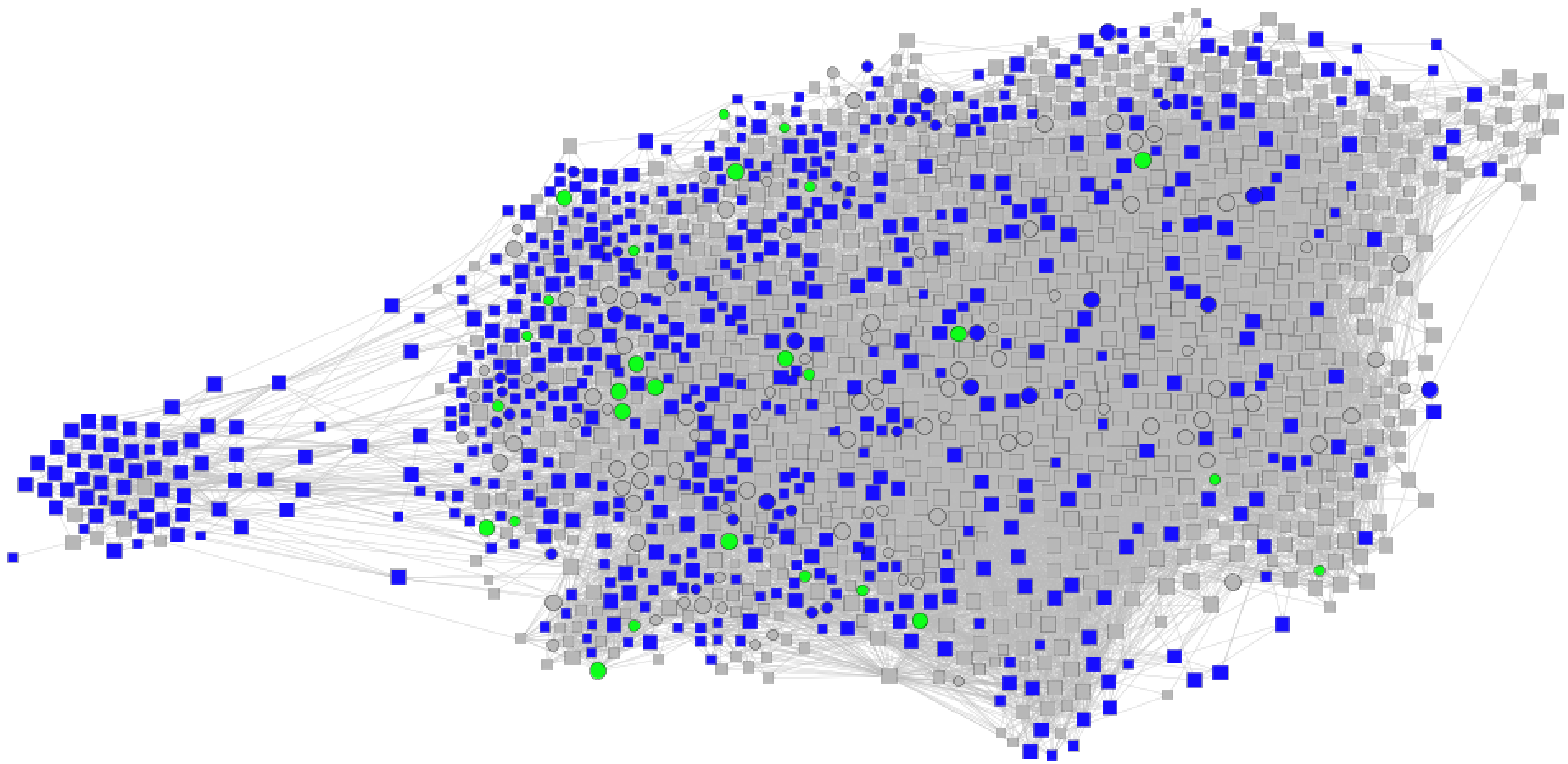
*Note: We model the Russian Product Space as a network in which nodes represent products exported by Russia with an RCA* $> 1$*, and links signify catalytic relations as represented in the Global Product Space. On this network, we apply the ACS detection algorithm to model the Russian Autocatalytic Set. Blue = Non-green products, Green = Green products, Gray = Products exported with an RCA* $< 1$

Figure 12: Indian Product Space and Autocatalytic Set



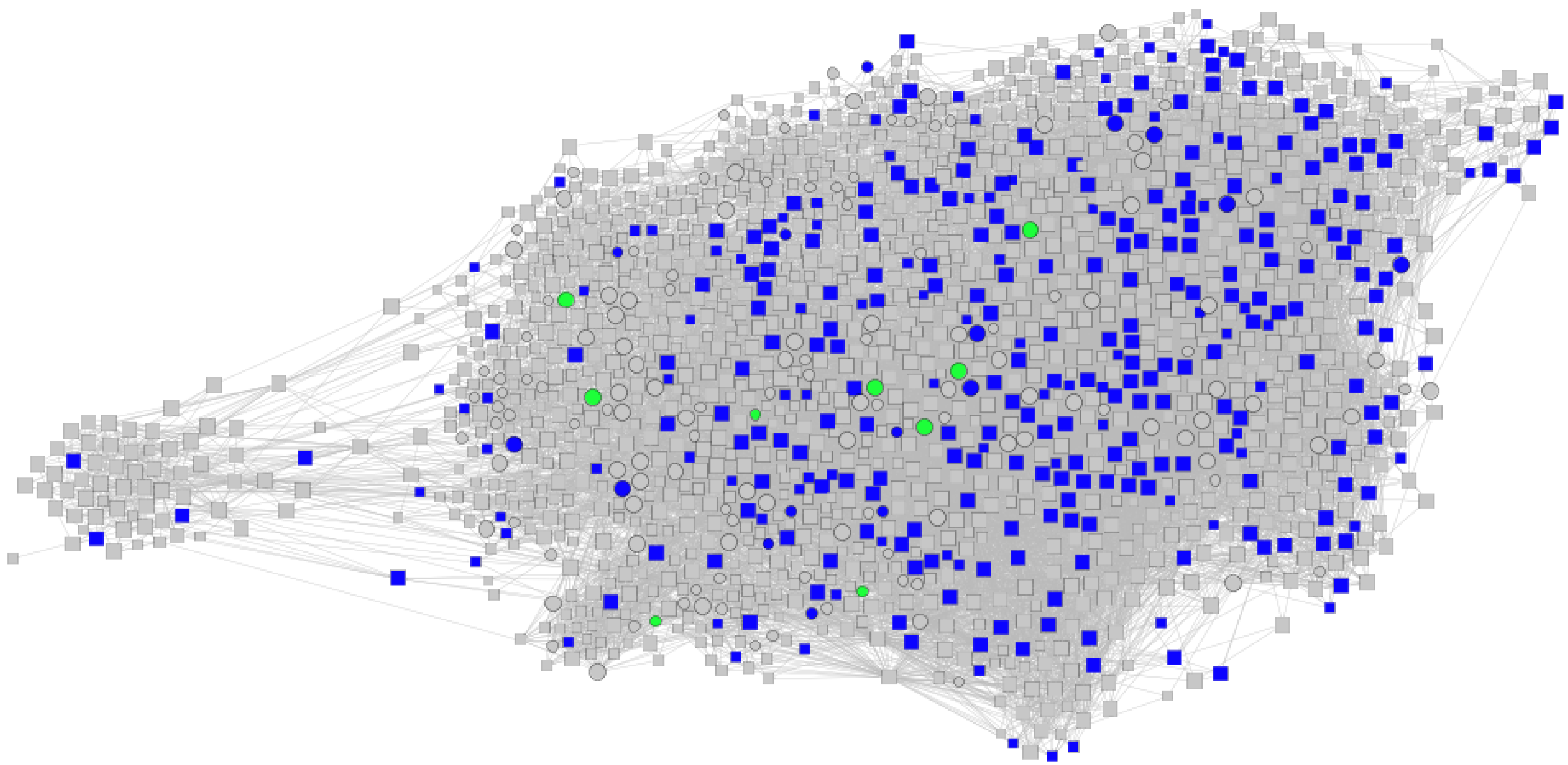
*Note: We model the Indian Product Space as a network in which nodes represent products exported by India with an $RCA > 1$, and links signify catalytic relations as represented in the Global Product Space. On this network, we apply the ACS detection algorithm to model the Indian Autocatalytic Set. Blue = Non-green products, Green = Green products, Gray = Products exported with an $RCA < 1$.*

Figure 13: Chinese Product Space and Autocatalytic Set



*Note: We model the Chinese Product Space as a network in which nodes represent products exported by China with an $RCA > 1$, and links signify catalytic relations as represented in the Global Product Space. On this network, we apply the ACS detection algorithm to model the Chinese Autocatalytic Set. Blue = Non-green products, Green = Green products, Gray = Products exported with an $RCA < 1$.*

Figure 14: South African Product Space and Autocatalytic Set



*Note: We model the South African Product Space as a network in which nodes represent products exported by South Africa with an RCA $> 1$, and links signify catalytic relations as represented in the Global Product Space. On this network, we apply the ACS detection algorithm to model the South African Autocatalytic Set. Blue = Non-green products, Green = Green products, Gray = Products exported with an RCA $< 1$.*

## DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2023
## *GREDEG Working Papers Released in 2023*

**2023-01** Thierry Blayac, Patrice Bougette & Florent Laroche
*What Drive HSR' Prices and Frequencies? An Analysis of Intermodal Competition and Multiproduct Incumbent's Strategies in the French Market*

**2023-02** Guilhem Lecouteux & Léonard Moulin
*Cycling in the Aftermath of COVID-19: An Empirical Estimation of the Social Dynamics of Bicycle Adoption in Paris*

**2023-03** Frédéric Marty
*Les politiques publiques européennes en faveur d'un cloud souverain : fondements, modalités de mise en oeuvre et évaluation critique*

**2023-04** Gérard Mondello
*"The Problem of the Social Cost" and the Legal Rights Question*

**2023-05** Sofia Patsali, Michele Pezzoni & Jackie Krafft
*Healthcare Procurement and Firm Innovation: Evidence from AI-powered Equipment*

**2023-06** Frédéric Marty
*De la maximisation de la valeur actionnariale à la prise en compte des intérêts des parties prenantes : les enjeux en termes de gouvernance d'entreprise*

**2023-08** Luca Fontanelli
*Theories of Market Selection: A Survey*

**2023-09** Paolo Zeppini & Jeroen C.J.M. van den Bergh
*Does COVID-19 Help or Harm the Climate? Modelling Long-run Emissions under Climate and Stimulus Policies*

**2023-10** David Laidler
*Peter Howitt – A Keynesian Still in Recovery*

**2023-11** Frédéric Marty
*Economie de la donnée: écosystèmes numériques, algorithmes et intelligence artificielle*

**2023-12** Giuseppe Attanasi, Anna Maffioletti, Giulia Papini, Patrizia Sbriglia & Maria Luigia Signore
*Beyond the Threshold: How Electoral Size-Dependent Uncertainty Affects Majority Determination*

**2023-13** Ivàn José Barreda Tarrazona, Agnès Festré & Stein Østbye
*Social Capital: Experimental Validation of Survey Measures*

**2023-14** Leonardo Ciambezi, Mattia Guerini, Mauro Napoletano & Andrea Roventini
*Rethinking Inflation in an Agent-Based Macroeconomic Model*

**2023-15** Giuseppe Attanasi, Barbara Buljat Raymond, Agnès Festré & Andrea Guido
*Augmented Reality Technology as a Tool for Promoting Pro-environmental Behavior and Attitudes*

**2023-16** Arnaud Persenda & Alexandre Ruiz
*Autocatalytic Networks and the Green Economy*