



# Anatomy of Green Specialisation: Evidence from EU Production Data, 1995–2015

Filippo Bontadini<sup>4,5</sup> · Francesco Vona<sup>1,2,3</sup>

Accepted: 26 May 2023 / Published online: 25 June 2023  
© The Author(s) 2023

## Abstract

We study green specialisation across EU countries and detailed 4-digit industrial sectors over the period of 1995–2015 by harmonizing product-level data (PRODCOM). We propose a new list of green goods that refines lists proposed by international organizations by excluding goods with double usage. Our analysis reveals important structural characteristics of green specialisation in the manufacturing sector. First, green production is highly concentrated, with 13 out of 119 4-digit industries, which are high-tech and account for nearly 95% of the total. Second, green and polluting productions do not occur in the same sectors, and countries specialise in either green or brown sectors. Third, our econometric analysis identifies three key features as relevant for green specialisation: (i) first-mover advantage and high persistence of green specialisation, (ii) complementarity with non-green capabilities and (iii) the degree of diversification of green capabilities.

**Keywords** Green goods · Green specialisation · Environmental policy stringency · Complementarity · Path dependency

**JEL Classification** Q55 · L60 · O44

## 1 Introduction

This paper provides new evidence of the production and specialisation of environmentally friendly goods across manufacturing sectors and European countries over the period of 1995–2015. Understanding the evolution and drivers of comparative advantage in green

---

✉ Francesco Vona  
francesco.vona@unimi.it

Filippo Bontadini  
fbontadini@luiss.it

<sup>1</sup> Department of Environmental Science and Policy, University of Milan, Milan, Italy

<sup>2</sup> Fondazione Eni Enrico Matteri (FEEM), Milan, Italy

<sup>3</sup> OFCE Sciences-Po, Paris, France

<sup>4</sup> LUISS Guido Carli University, Rome, Italy

<sup>5</sup> SPRU, University of Sussex, Falmer, UK

productions is particularly important given the growing interest around the so-called green economy as a way to create new business opportunities related to environmental preservation and climate change mitigation. This view recently culminated in the launch of the European Green Deal by the European Commission in response to the COVID-19 pandemic (Agrawala et al. 2020; Chen et al. 2020). Developing a first-mover advantage in new high-demand products such as electric cars and PV panels was also a strategic goal of the generous fiscal stimulus implemented by President Obama after the great recession, the so-called American Recovery and Reinvestment Act (Popp et al. 2021).

Despite its key strategic role for a country's future competitiveness (Fankhauser et al. 2013), data constraints have so far limited empirical research on the green economy. We contribute to filling this gap by examining the drivers of green specialisation in EU manufacturing industries over two decades. In doing so, we assemble a new dataset that leverages granular product-level data to construct measures of green production shares and Balassa indexes of specialisation for 4-digit manufacturing industries. Understanding where green productions are located and why, is a first step to assess the potential benefits of a demand push for the green economy. For instance, job creation effects depend on where production is located rather than on where knowledge is created.

This paper contributes to the literature on industry specialisation being the first to systematically focus on green industries, namely on industries producing goods that reduce negative environmental impacts. In our main empirical analysis, we build on the standard framework used to study the drivers of industry specialisation (Midelfart-Knarvik et al. 2000; Romalis 2004; Bombardini et al. 2012; Nicoletti et al. 2020). Mulatu et al. (2010) adopt this framework to study the role of environmental policies and other structural drivers in the location choices of polluting industries in a static, cross-sectional analysis. We enrich this framework by exploiting the panel dimension of our data. This is particularly relevant to study patterns of comparative advantage in emerging and new products such as green ones.

Moreover, we take stock from the findings of the recent empirical literature on green innovation to include three new drivers of green specialisation: i. path dependency and persistency of first-mover advantage (e.g., Aghion et al. 2016); ii. complementarity with non-green capabilities (e.g., Perruchas et al. 2020); and iii. diversification of the knowledge base (e.g., Colombelli and Quatraro 2019). With respect to the few papers on green specialisation using export (Mealy and Teytelboym 2020) or patent data (Perruchas et al. 2020; Barbieri et al. 2020), we highlight the importance of observing the fine-grained structure of production across sectors and countries (and not only across countries) to understand specialisation patterns. In doing so, we provide a new angle and a new dataset for the analysis of green innovation that so far has mostly used patents (e.g., Popp 2002; Nesta et al. 2014; Cabel and Dechezleprêtre 2016) or self-reported measures of innovation (Frondel et al. 2007, Horbach et al. 2012). While most climate (e.g. Nordhaus and Boyer 2000) and endogenous growth models (e.g. Bovenberg and Smulders 1995) give prominent importance to R&D-driven green innovation, knowledge is also created during the production stage, through learning by doing, as postulated in growth theory (Romer 1990) and innovation studies (Arrow 1971; Hatch and Mowery 1998; Clarke 2006). Overall, patents and productions capture two complementary channels of the process of building a green comparative advantage. We offer new data and evidence on the production channel, which has been comparatively understudied in the literature.

A crucial step to provide new evidence on green specialisation is to construct a time-consistent measure of green production that varies at the country-year-sector (detailed 4-digit NACE rev. 2 sectors) level. To this end, we harmonize a product-level dataset

compiled by Eurostat for the manufacturing sector, called PRODCOM, using the methodology proposed by Van Beveren et al. (2012). To identify green products, we first select candidate lists of products that have been proposed during recent international negotiations at the World Trade Organization (WTO), as well as by the OECD. We then refine these lists by eliminating goods with double usages to obtain a favourite list of products that we use to construct measures of green production at the sector-by-country level. While previous works used product-level data at the country level to study trade patterns in green products (Fraccascia et al. 2018; Mealy and Teytelboym 2020), our new dataset is the first that allows to study green production (rather than just trade) along three dimensions: country, year and highly granular 4-digit industries.

Our empirical analysis uncovers a set of novel facts concerning the structure of green production in European manufacturing industries.

First, we show that green production is just 2–2.5% of total manufacturing production in Europe. This is in the ballpark of the most precise US measures of the green economy (Elliott and Lindley 2017; Vona et al. 2019). This finding suggests that a very granular dataset such as the one assembled in this paper is crucial to studying green specialisation.

This insight is further supported by our second key finding: green production is extremely concentrated in a set of high-medium tech industries mostly producing capital goods. At the 2-digit level, 9 out of 26 manufacturing industries have non-null green production. However, of the 119 4-digit industries contained in those green 2-digit industries, only 21 are green, and 13 of those represent 94.9% of the total green production.

Last, we find that, even at a rather coarse level of aggregation (2-digit industries), polluting and green productions occur in two disjoint sets of industries. Consequently, the industries bearing the cost of environmental regulation will be different from those that will receive most of the benefits of the green industrial policies. This difference in the industry exposure to environmental policies also has consequences at the country level. Northern countries, especially Denmark, Sweden and Germany (along with Austria), exhibit a persistent green comparative advantage. In contrast, lower income countries, such as Greece, Romania and Bulgaria and some traditionally industrial economies, such as Italy and Belgium, have retained a specialisation in polluting productions.

Going beyond these descriptive facts, we examine the drivers of green specialisation at the country-industry level using panel data econometrics, following the approach described above (e.g., Midelfart-Knarvik et al. 2000; Romalis 2004). Consistent with the descriptive evidence, green specialisation exhibits path dependency, confirming the importance of first-mover advantages. Although our empirical strategy is not designed for policy evaluation, environmental policies seem to contribute to building a first-mover advantage in green industries, while they do not appear to be effective for catching-up or maintaining such advantage. Our regressions also reveal a complementarity between green and non-green specialisation within the same narrowly defined 4-digit industries, although the magnitude of the association is smaller than the persistency of the lead start advantage. Finally, diversifying the portfolio of green products with comparative advantage is also important for sustaining green specialisation.

The remainder of the paper is structured as follows. Section 2 discusses conceptual issues in defining green production, provides details on the PRODCOM data we use to compile our dataset and on how we refine existing lists to obtain our favourite one. Section 3 presents descriptive evidence on industry-level dynamics of green production across European countries. Section 4 looks at specialisation in green production at the country level and relates it to environmental policies. Section 5 builds on the descriptive results of the previous sections and looks at drivers of country-industry specialisation in green

specialisation. In Sect. 6, we summarize our main findings and provide future research avenues.

## 2 A New Measure of Green Production

This section is organized as follows. In Sect. 2.1, we discuss the conceptual issues in measuring green production. In Sect. 2.2, we present our main source of data, PRODCOM and we illustrate how to use PRODCOM to measure green production. Finally, Sect. 2.3 validates our favourite list of green products—which we refer to as the PRODCOM list henceforth—against other lists.

### 2.1 Conceptual Issues

The definition of green production presents several conceptual challenges related to the understanding of what “green” means and how such definitions can be operationalized in the data. We should emphasize from the outset that we consider “green” referring to the relationship between economic activities and the environment, leaving aside other dimensions related for instance to whether the green economy should encompass social inclusion and equality of opportunities (e.g., Merino-Saum et al. 2020).

The first conceptual issue is whether we consider an activity (i.e., a product or a service) green in terms of the effective pollution content of its production (*process approach*) or in terms of its potential to minimize the harmful impacts of production on the environment (*output approach*). The first approach is intuitive: it uses direct and indirect pollution generated in producing a good as measure of the inverse of the product greenness. However, data limitations make it difficult to devise a measure of the pollution content of products that varies across countries and years (Sato 2014). While input–output methodology has been used to better assess the environmental footprint of production, unfortunately available input–output tables only include a limited number of countries, years and highly aggregated sectors, yielding mixed and incomplete results on the pollution content of different productions (Rodrigues et al. 2018).

The output approach emphasizes the potential of certain products to be beneficial for the environment, and it is the preferred approach for defining most lists of green products or activities. For instance, both the Green Goods and Services Survey (GGS) of the Bureau of Labor Statistics in the US (e.g., Elliott and Lindley 2017) and the Eurostat definitions of green products (Eurostat 2016) use an output-based approach. To illustrate the difference between these two approaches, one can consider wind turbines: even though they fulfil an unequivocally green function, the process, emission-based, approach would not consider them very green due to the high pollution intensity of the iron that is necessary for their production.

In this work, we focus on the output approach. In line with this choice, our main conceptual challenge is identifying which functions are particularly beneficial to the environment. This is far from straightforward: products fulfil functions that differ in their potential for reducing pollution based on their underlying technology, such as end-of-pipe and integrated technologies (Frondel et al. 2007).<sup>1</sup> A crucial conceptual issue here is that the same

<sup>1</sup> End-of-pipe technologies limit pollution from production processes without changing these processes in essence (e.g., waste-water treatment, catalytic converters or exhaust-gas cleaning equipment). Integrated

product can have different usages and thus different environmental impacts. For example, pipes and water tanks may be considered green when used for water and waste management purposes, but they will not be green when used for other activities (Steenblik 2005), such as textile production that involves intensive water consumption. Altogether, these issues make it difficult to find a widely accepted conceptual definition of what a green product is. Operationalizing a definition of green products is even more difficult because standard statistical classifications are not designed to identify green products (Steenblik 2005; Sauvage 2014). This increases the likelihood that a green products' list contains false negatives (products that are green but are excluded from the list) and false positives (products that are not green but that are nonetheless included). This paper proposes to mitigate the data shortcomings and conceptual ambiguities discussed above using a new dataset, PRODCOM, where product codes and descriptions are available at a highly disaggregated level.

## 2.2 Measuring Green Production Using the PRODCOM Data

### 2.2.1 The Dataset

In the PRODCOM dataset, Eurostat collects very detailed information on manufacturing production values in Europe, covering on average, 4288 single products per year. The dataset is available for the years between 1995 and 2015 for the core European countries, while detailed data on production in Eastern European countries has been collected from 2001 onwards.<sup>2</sup>

To identifying green production across countries and industries, PRODCOM presents two advantages. First, the data are easily linkable to existing lists of green products. Second, the PRODCOM classification is nested within the European industrial classification NACE: each PRODCOM code has eight digits, the first four of which correspond to NACE industry codes. This feature allows assigning each product to a 4- (and 2-digit) industry and computing the industry's share of green production.

The use of PRODCOM presents some practical issues, which are extensively discussed in Online Appendix A.1. Notably, PRODCOM codes are updated yearly making it difficult to build a consistent panel of products. We deal with this issue using the methodology developed by Van Beveren et al. (2012) (VBBV henceforth) to harmonize the PRODCOM data over time. In a nutshell, the VBBV methodology identifies chains of product codes, which change over time due to statistical reclassification, and attributes a "synthetic code" to each chain that does not change over time, thus obtaining a consistent measure over time by product and sector. Combined with any list of green products, the VBBV methodology allows us to classify a synthetic code as either green (*g*) or not green (*ng*) and then to allocate these products to 4-digit NACE rev. 2 industries. For each industry, we compute the share of green production as follows:

---

Footnote 1 (continued)

technologies prevent pollution at the source, replacing less clean technologies: wind turbines are a clear example of this kind of product.

<sup>2</sup> Countries for which data from 1995 on is available include: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Portugal, Spain, Sweden and the United Kingdom. From 2001 on, our data include: Bulgaria, Croatia, Estonia, Hungary, Latvia, Lithuania, Romania, Slovakia and Slovenia. Poland is included from 2003 onwards.

$$Green\ Share_{ijt} = \frac{\sum_g y_{ijt,g}}{\sum_g y_{ijt,g} + \sum_{ng} y_{ijt,ng}} \quad (1)$$

where we divide the production of green goods in country  $i$ , industry  $j$  at time  $t$ , by the sum of both green and non-green production in the same country-industry-year combination.

## 2.2.2 Defining a Favourite list of Green Products

A key step of our analysis is to define a favourite list of green goods to implement the VBBV harmonization procedure. Historically, various lists of green products emerged as part of international negotiations to reduce the tariffs on a set of goods that are crucial for low-carbon transitions and sustainable development in general (WTO 2001; APEC 2012). The rationale for such negotiations is that decreasing tariffs on green products favours their diffusion and thus reduces abatement costs (World Bank 2007; Hufbauer and Kim 2010), especially in developing countries (Dutz and Sharma 2012; World Bank 2012).

Unfortunately, in pursuing this important goal, political economy considerations added a source of ambiguity to the definition of what is green. Indeed, each country negotiates “green” tariff reductions on the goods for which they have a comparative advantage rather than on truly green goods (Balineau and de Melo 2011; de Melo and Solleder 2018). The resulting disagreement on a final list of green goods was one of the reasons for halting the negotiations on trade in environmental goods in 2016 (European Commission 2019).<sup>3</sup>

Among the lists of green products proposed in such negotiations, the most comprehensive is the Combined List of Environmental Goods (CLEG) of the OECD, which encompasses three lists: the Plurilateral Environmental Goods and Services (PEGS) list developed by the OECD itself, the list proposed by the Asian Pacific Economic Cooperation (APEC) forum and the list defined by the so-called WTO Friends group. These lists are compiled using the Harmonized System (HS), the most widely used product classification system for trade across countries that can be linked to PRODCOM codes using crosswalks provided by Eurostat.

Additionally, although there is no official list of green products compiled by Eurostat, the list of the German Statistical Office follows the Eurostat criteria to define environmental goods.<sup>4</sup> We consider the union of the CLEG and German lists to provide a list of “potential green goods” that consists of 902 products. We refine this broad list to reach our favourite PRODCOM green list by excluding goods with multiple usages. In doing so, we review the product descriptions of the PRODCOM codes and manually exclude products with both green and non-green usages, such as tanks, industrial ovens, baskets, and mats. Among the goods with double usages, we retain only those related to the monitoring and analysis of environmental variables such as thermostats and apparatus equipment for meteorology and the chemical analysis of water. These products are included in all three lists composing the CLEG list, indicating a consensus around their green potential.

<sup>3</sup> For instance, bicycles have been at the centre of controversy between China and the European Union.

<sup>4</sup> Environment protection activities “have as their main purpose the prevention, reduction and elimination of pollution and of any other degradation of the environment” and resource management, i.e., the “preservation, maintenance and enhancement of the stock of natural resources and therefore the safeguarding of those resources against depletion” (Eurostat 2016, p.15). This narrow definition excludes products that do not match any criterion but that reduce environmental impacts in other sectors, such as LEDs.

Our cleaning procedure leaves us with 221 (from 4288 products included in the PRODCOM data and 902 products from the union of the CLEG and German lists) green products.

### 2.2.3 Advantages of Production Data

The other lists we use, in the next subsection, for sake of comparison are largely based on trade data, while we rely on production data. It is therefore important to highlight the key advantages of this choice. We focus here on approaches that have relied on the use of secondary data, rather than the collection of original data through surveys.

First, a vast literature uses trade data and a variety of existing lists of green products to study trade patterns (He et al. 2015; Cantore and Cheng 2018; Fraccascia et al. 2018; Tamini and Sorgho 2018; Mealy and Teytelboym 2020) and their effects on emission reduction (Zugravu-Soilita 2018, 2019). We also rely on such lists to build our own list, as just discussed. However, trade represents only a small portion of an economy and exporting firms are a non-random sample of large and highly productive firms (Melitz 2003; Bernard et al. 2007, 2012). As a result, using data on total production, rather than just the subsection of production that is exported, is likely to provide a more accurate picture of how green production is distributed across countries and industries.

Second, another well-established strand of literature has relied on data on patenting activity (Jaffe and Palmer 1997; Popp 2002; Nesta et al. 2014; Caelal and Dechezleprêtre 2016; Sbardella et al. 2018; Perruchas et al. 2020). A key advantage of using patents is that patent classification explicitly identifies green patents—e.g., the tag Y02 provided by the European Patent Office (EPO). However, patent data refers to where knowledge is created, but not so much on where production actually takes place and where green jobs are created (Vona et al. 2019). Moreover, and crucially, patent data only captures codified knowledge, while the literature on innovation studies has shown that other non-codified ways of learning are also crucial to economic activity (Cowan et al. 2000; Johnson et al. 2002; Balconi et al. 2007). In this respect, PRODCOM data provides a reliable output measure of the green economy able to capture such learning effects.

Third, an interesting approach is that of Shapira et al. (2014), who develop a set of keywords that identify green products and then use these terms to identify green firms based on their reported business description in a database compiled by Bureau van Dijk for the UK. Retrieving green production using firm-level data is very promising but such data are usually available only for a selected sample of firms and not for several countries. In contrast, the PRODCOM dataset is based on administrative sources of data, offering reliable statistics across industries, years and countries.

Finally, it is also possible to use a combination of trade, patent, and production data. Frankhauser et al. (2013) propose such an approach to identify potential winners of the “green race”. They combine a wide range of sources 110 industries, across 8 countries over only 2 years (2005–2007). Using PRODCOM data, we have information across most European countries, over a much longer period (1995–2015) and we identify green products starting from over 4000 products rather than 110 industries. Moreover, our measure of green production is continuous, rather than binary, providing a more nuanced picture of green activities across countries and sectors.

Note that all the approaches above only look at the manufacturing sector, leaving service industries aside. Unfortunately, PRODCOM data offers no remedy to this limitation as it only covers the manufacturing industry. Note also that PRODCOM data only covers European countries, while other works using patent and trade data, cover a broader group



of countries. Although this is a drawback of PRODCOM data, Europe represents an important share of global green production and observing green production at such granular level of industry-country-year aggregation more than compensates for this shortcoming.

### 2.3 Comparisons with Other Lists of Green Products

While it is not possible to prove unequivocally that our list is the most adequate to identify green products, the comparison with other lists allows to highlight some key advantages. We compare here our favourite PRODCOM list with five broader lists (CLEG, German list, APEC, PEG and WTO2009) and two narrower lists (WTO Core and Core CLEG). We discuss each of these lists in greater details in Table A.1 of the Online Appendix A.2.

In Table 1 we correlate vectors of dummy variables indicating the presence of a certain product in each list. While the correlation across broader lists (PEGS, APEC, WTO2009 and CLEG) is quite high, narrower lists, such as the WTO Core and Core CLEG lists, are weakly correlated with each other. For instance, the WTO Core and Core CLEG lists share only one green product, i.e., spectrometers using optical radiation. Our favourite PRODCOM list exhibits a quite strong correlation with the WTO2009 list (with a correlation coefficient of 0.49), as well as its narrow version, the WTO Core list (0.3) and with the PEGS list (0.58). We also find a high correlation coefficient (0.45) between our PRODCOM list and a “Core list”, which is defined as the union of the WTO Core and Core CLEG lists. This implies that our favourite PRODCOM list identifies a large portion of products that are included in either of the two most restrictive lists, reassuring us on the credibility of our favourite list. To give a few examples, these products include end-of-pipe technologies such as machinery for purifying gases and liquids as well as integrated technologies such as solar cells and monitoring equipment for physical and chemical analysis.

Figure 1 shows the overlap between our favourite PRODCOM list, the broadest CLEG list, German list, and the narrowest Core list. We find that 79 out of 147 products from the German list that are not included in any other list and that the CLEG has several products, 512 out of 819, that are not part of other lists. Such products include multi-usage products such as tanks, industrial ovens and machinery for sorting and grinding material.

The narrow Core list is fully contained in the CLEG list, but it also shares products with the German list and our favourite PRODCOM list. This suggests that there is a consensus around products included in the Core list, but we find that important green products are not included in the Core list. Indeed, the Core list focuses on products whose function is to directly combat pollution with end-of-pipe technologies (i.e., water and waste management equipment) rather than on key integrated technologies (i.e., wind turbines). This list also leaves out secondary environmental products that offer more environmentally sustainable mobility options—such as bicycles—and environmental monitoring equipment.<sup>5</sup>

In conclusion, our favourite PRODCOM list seems more accurate than other available lists. On the one hand, broader lists, such as the CLEG, German, and APEC lists, include products with multiple non-green usages. On the other hand, narrower lists leave out

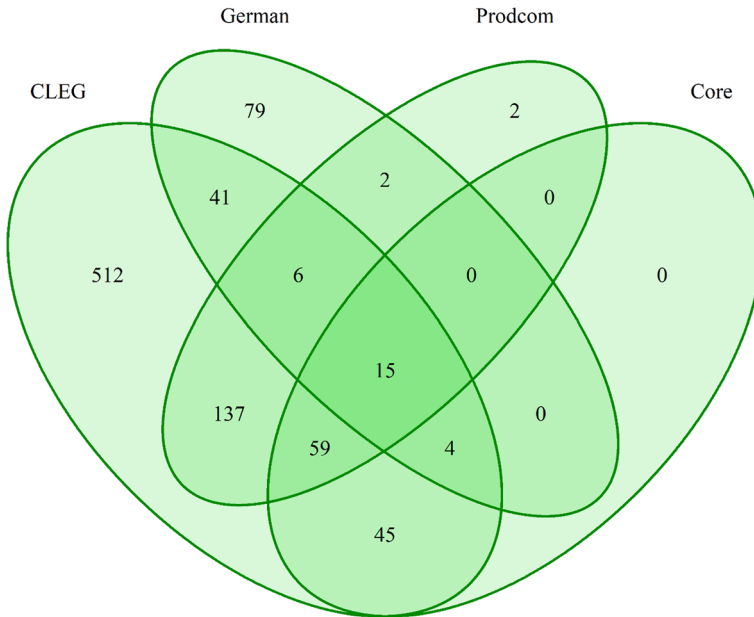
<sup>5</sup> Gas turbines are included in the WTO list and so are also part of our PRODCOM list. Clearly, their treatment is problematic. On the one hand, they are a transition technology, so they can be considered green. On the other hand, they produce GHG emissions, so they are brown. In conclusion, gas turbines' output does not have a direct positive effect on the environment. In fact, burning natural gas produces GHG and methane emissions and does not reduce pollution in other industries. Hence, there are no reasons to consider gas turbines as green, neither from the output nor from the process approach.



**Table 1** Correlation table among green product lists

	(1) CLEG	(2) WTO 2009	(3) PEGS	(4) <i>PRODCOM</i> ( <i>favourite</i> )	(5) APEC	(6) German list	(7) Core (WTO+CLEG)	(8) WTO Core	(9) CLEG Core
CLEG	1								
WTO 2009	0.84**	1							
PEGS	0.73**	0.47**	1						
<i>PRODCOM</i> ( <i>favourite</i> )	0.49**	0.31**	0.58**	1					
APEC	0.46**	0.49**	0.41**	0.31**	1				
German list	0.16**	0.15**	0.14**	0.12**	0.17**	1			
Core (WTO+CLEG)	0.37**	0.37**	0.35**	0.45**	0.44**	0.13**	1		
WTO Core	0.29**	0.25**	0.27**	0.3**	0.16**	0.04**	0.77**	1	
CLEG Core	0.23**	0.28**	0.24**	0.35**	0.51**	0.16**	0.65**	0.03**	1
Number of goods	819	604	470	221	206	147	123	78	47

Authors' own calculation on PRODCOM data. The table reports correlation coefficients of dummy variables indicating the presence of a certain product in a given list across different lists. The last row reports the number of PRODCOM product codes within each green product list. For further details about the lists of green goods, see Online Appendix A. \* $p < 0.05$  \*\* $p < 0.01$



**Fig. 1** Overlap of PRODCOM product codes among selected lists of green goods. *Notes* Authors' elaboration on PRODCOM data. The figure depicts the overlap among four existing lists of green goods, the numbers represent the number of PRODCOM product code that fall within each category. For further details about the lists of green goods see Online Appendix A

integrated technologies such as wind turbines, electric cars, and environmental monitoring equipment. Our favourite PRODCOM list strikes a balance between these two extremes by focusing on single-usage products and by including both products that directly affect the environment and products that reduce pollution and energy use in other industries (such as LED bulbs, heat pumps and batteries). In section D of the Online Appendix, we replicate our main results using the CLEG list. We choose the CLEG list as term of comparison because it is a well-established and broad list that includes several multi-usage products. To be clear, our aim is to identify a list of core green products that are with no doubt the basis of green specialisation, not to argue that multiple-usage products have no role to play in achieving the transition to a greener economy. In general, our results using the CLEG list led to estimates of the share of the green economy that are well above other benchmarks existing in the literature (for the US, e.g., Elliott and Lindley 2017; Vona et al. 2019), further validating our choice of a favourite list that excludes them.

### 3 Green Production Across Industries

We begin by exploring the industry dimension of the data using the share of green production relative to total production as key statistics. Using such measure allows us to capture the high degree of heterogeneity in green production across and within industries.

**Table 2** Green and polluting production by 2-digit industries

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NACE	Label	Mean green share 2010	Mean green share 2005	Mean green share 2010	Share of total green production	Absolute Change 2005–2015	Average GHG intensity
28	Manufacture of machinery and equipment n.e.c	0.074 (0.068)	0.084 (0.083)	0.096 (0.098)	0.28	0.022	0.54
27	Manufacture of electrical equipment	0.108 (0.166)	0.103 (0.078)	0.162 (0.217)	0.22	0.054	0.30
26	Manufacture of computer, electronic and optical products	0.069 (0.06)	0.121 (0.131)	0.103 (0.076)	0.22	0.034	0.30
30	Manufacture of other transport equipment	0.281 (0.292)	0.346 (0.318)	0.38 (0.334)	0.13	0.098	0.61
33	Repair and installation of machinery and equipment	0.022 (0.031)	0.033 (0.024)	0.028 (0.026)	0.04	0.006	0.74
29	Manufacture of motor vehicles, trailers and semi-trailers	0.002 (0.01)	0.007 (0.031)	0.003 (0.011)	0.01	0.001	0.61
31	Furniture	0 (0)	0 (0)	0 (0)	0	0	0.74
32	Other manufacturing	0 (0)	0 (0)	0 (0)	0	0	0.74
16	Products of wood, cork, straw, plaiting	0 (0)	0 (0)	0 (0)	0	0	0.88
22	Rubber and plastic products	0 (0)	0 (0)	0 (0)	0	0	0.94
13	Textiles	0 (0)	0 (0)	0 (0)	0	0	0.97
14	Wearing apparel	0 (0)	0 (0)	0 (0)	0	0	0.97
15	Leather and related products	0 (0)	0 (0)	0 (0)	0	0	0.97
17	Paper and paper products	0 (0)	0 (0)	0 (0)	0	0	1.18
18	Printing and reproduction of recorded media	0 (0)	0 (0)	0 (0)	0	0	1.18
10	Food products	0 (0)	0 (0)	0 (0)	0	0	1.45
11	Beverages	0 (0)	0 (0)	0 (0)	0	0	1.45
12	Tobacco products	0 (0)	0 (0)	0 (0)	0	0	1.45

Table 2 (continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NACE	Label	Mean green share 2005	Mean green share 2010	Mean green share 2015	Share of total green production	Absolute Change 2005–2015	Average GHG intensity
<i>Polluting industries</i>							
19	Coke and refined petroleum products	0 (0)	0 (0)	0 (0)	0		44.99
23	Other non-metallic mineral products	0.029 (0.029)	0.033 (0.022)	0.033 (0.026)	0.05	0.003	7.78
20	Chemicals and chemical products	0 (0)	0 (0)	0 (0)	0	0	5.11
21	Basic pharma. products, preparations	0 (0)	0 (0)	0 (0)	0	0	5.11
25	Fabricated metal products, exc. machinery	0.018 (0.018)	0.019 (0.016)	0.017 (0.014)	0.05	-0.001	4.23
24	Basic metals	0.006 (0.021)	0.007 (0.023)	0.008 (0.03)	0.01	0.002	4.23

Authors' elaboration on PRODCOM data. Production values are deflated to have data at constant prices, with 2010 as base year. The definition of green products used here is explained in Sect. 2 (PRODCOM in Fig. 1). Columns 3–5 report the mean green share of production with the standard deviation in brackets of each industry for the years 2005, 2010 and 2015, respectively. Coke and refined petroleum products is not included in PRODCOM until 2005, as PRODCOM coverage is not stable over time and doesn't include fuel related products. Column 6 reports the share that green production of each industry represents in total green production. Absolute changes 2005–2015 refer to industries' average green shares of production. Polluting industries are identified as the 5 industries with the highest average GHG intensity computed with WIOD, for further detail see Online Appendix B

### 3.1 Aggregated Industries: Green Versus Brown Production

In Table 2, we explore the variability in the share of green production across 2-digits industries. This higher level of aggregation allows to compare the output-based and process- (emission-) based definitions of green production. We report the mean and standard deviation of green shares for each industry, as well as the average GHG intensity. As mentioned in Sect. 2.2, the number of countries included in the PRODCOM data is unbalanced, thus we focus on 2005, 2010 and 2015, where we have a balanced panel of countries.

We find that green production is highly concentrated in a few industries. While most 2-digit sectors (17 out of 26) have no production of green goods, four industries emerge as the key players in the green transition: i. Computer, electronic and optical equipment, which includes photovoltaic panels; ii. Electrical equipment, which includes equipment for the control and distribution of electricity; iii. Machinery and equipment, which includes wind turbines; and iv. Other transport equipment, which includes railway stocks. Remarkably, these four industries represent 85% of the total green production (column 6). Within these four industries, we also observe a rather high coefficient of variation (standard deviation), which indicates a high degree of heterogeneity in green production across countries. Over time, average green shares increase in all the four greenest industries, in contrast with the stability of the average green share in other green industries.

Importantly, the four industries with a high green potential have a few other characteristics that are of strategic interest for industrial policy. First, they are all high- or medium-high tech industries (Eurostat 2015; Galindo-Rueda and Verger 2016—see also in Online Appendix E for a list of high-tech industries<sup>6</sup>) that have large job multipliers in local labour markets (Moretti 2010; Vona et al. 2019) and are conducive to economic growth (McMillan et al. 2014; Szirmai and Verspagen 2015). Second, specialisation in these sectors requires a strong pool of pre-existing capabilities (Hidalgo et al. 2007; Mealy and Teytelboym 2020), particularly engineering and technical skills that are prevalent in green jobs (Vona et al. 2018).

To compare the output-based and the process-based definition of green production, the last column of Table 2 reports greenhouse gas (GHG) intensity for the same 2-digit manufacturing industries. We rely on the environmental accounts of the World Input–Output Database (WIOD) that include the energy and GHG content of domestic production of each 2-digit industry for 15 countries between 1995 and 2009. We compute GHG (CO<sub>2</sub>, N<sub>2</sub>O and CH<sub>4</sub>, aggregated according to their global warming potential) intensity as the sum of direct and indirect emissions per unit of value added from each industry, country, and year. A well-known cluster of brown<sup>7</sup> industries stands out in terms of total (direct and indirect) emissions (Wiebe and Yamano 2016; de Vries and Ferrarini 2017): coke and refined petroleum products, other non-metallic mineral products, chemicals and chemical products, basic pharmaceutical products and pharmaceutical preparations, and basic metals and the manufacturing of fabricated metal products, except machinery. In the remainder of

<sup>6</sup> The only exception is sector repair and installation of machinery and equipment (NACE rev. 2 industry 33).

<sup>7</sup> In the remainder of this paper we define brown (or polluting) industries as those that have high GHG intensity as described above. Green production in contrast is defined based on the “output” approach, also detailed above; conversely, non-green production refers to economic activities that are not green according to the output-based definition.

this paper, we treat the entire production process of these brown industries as polluting (see Online Appendix B and Marin and Vona 2019, for details).

Remarkably, comparing columns 3 to 5 with column 8 of Table 2, we observe no overlapping between green production and pollution intensity. This fact has two main implications. First, from a conceptual point of view, the process- and output-based approach capture different aspects of the green economy but are not in contradiction with each other and in fact end up identifying similar “green” manufacturing industries. Second, the two approaches are complementary for analysing policy impacts and understanding the distributional effects of environmental policies. While the competitiveness of brown industries is potentially harmed by an increase in environmental policy stringency (Dechezleprêtre and Sato 2017), green sectors benefit from the indirect demand for pollution abatement equipment, technical know-how and integrated technologies (Horbach et al. 2012; Vona et al. 2019).<sup>8</sup> In other words, the two well-known channels through which environmental policies affect competitiveness, namely, the cost channel (eventually leading to relocating polluting industries abroad, the pollution haven hypothesis) and the innovation channel (the so-called Porter hypothesis; Ambec et al. 2013), impact different sets of industries. Because this result is obtained at a quite aggregate industry level due to limited data availability on emission content of more disaggregated industries, the lack of overlapping between green and GHG-intensive industries may mask substantial heterogeneity at the product-level.<sup>9</sup>

### 3.2 Disaggregated Industries: Identifying High-Green-Potential Industries

We compare green and polluting production at 2-digit level of aggregation due to data constraints on sectoral GHG-intensity. However, the granularity of PRODCOM data allows us to compute the shares of green production for 4-digit industries. This is important to understand which industries green production is concentrated in.

Table 3 reports key statistics on a set of 4-digit industries where green production is greater than zero in at least one year. The main takeaway is again that green production is highly concentrated also at the 4-digit level. Of the 119 4-digit industries among the 2-digit industries with a green production greater than zero, only 21 are green. Moreover, we find that 11 out of these 21 industries have a maximum green production of 100% for at least one country and year.

After ranking industries by their average share of green production, we observe a first group of eight extremely green industries, from “bicycle and invalid carriage manufacturing” to “non-domestic cooling and ventilation equipment manufacturing”. For these

<sup>8</sup> When we use a broader set of green products, such as the CLEG list, we find that core green industries from Table 2 still rank among the top green industries. We also find that fabricated metal products, (exc. machinery) and other non-metallic mineral products also exhibit large green shares of production. This is the result of the CLEG list including multiple use products such as tanks, taps and plastic containers that can be used for water and waste management purposes but also have a high emission content. This is discussed more at length in Online Appendix D—see Table D.1. Despite these, our results confirm that green production is still heavily concentrated in few industries, with little overlap with polluting industries.

<sup>9</sup> Although at the aggregate level we do not observe any overlapping, it may be that, for some specific green products such as wind turbines, there can be a trade-off between the environmental impact of producing green products and the potential environmental benefits in their usage. Studying such trade-offs at the product level require more detailed data that, to the best of our knowledge, are only available for India (Barrows and Ollivier 2018) and it is well beyond the scope of this study.

**Table 3** Distribution of green production shares across green industries at 4 digits NACE

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NACE	Label	Mean	Median	Max	Standard deviation	Change 1995–2015	Change 2001–2015	Share of green production
<i>High green potential industries</i>								
3092	Manufacture of bicycles and invalid carriages	0.78	0.82	1	0.24	0.11	0.14	3.12
3020	Manufacture of railway locomotives and rolling stock	0.71	0.80	1	0.28	-0.08	-0.06	9.99
2530	Manufacture of steam generators, except central heating hot water boilers	0.55	0.54	1	0.35	0.21	0.34	1.91
2312	Shaping and processing of flat glass	0.4	0.34	1	0.30	0.04	0.07	4.83
2712	Manufacture of electricity distribution and control apparatus	0.39	0.34	1	0.23	0.03	0.03	16.86
2651	Manufacture of instruments and appliances for measuring, testing, etc	0.37	0.37	1	0.19	0.04	-0.04	18.29
2829	Manufacture of other general-purpose machinery n.e.c	0.29	0.24	1	0.22	0.12	0.16	7.75
2825	Manufacture of non-domestic cooling and ventilation equipment	0.28	0.28	1	0.18	-0.07	-0.04	11.18
2811	Manufacture of engines and turbines, except aircraft, vehicle and cycle engines	0.21	0.07	1	0.30	0.14	0.06	8.61
2611	Manufacture of electronic components	0.14	0.01	1	0.27	0.13	0.13	3.85
2740	Manufacture of electric lighting equipment	0.13	0.12	0.66	0.10	0.00	-0.04	3.48
2752	Manufacture of non-electric domestic appliances	0.11	0.03	0.50	0.14	-0.03	0.02	1.05
3320	Installation of industrial machinery and equipment	0.08	0.06	0.67	0.08	0.08	0.05	3.98
<i>Marginally green industries</i>								
2410	Manufacture of basic iron and steel and of ferro-alloys	0.04	0.00	1.00	0.18	0.03	-0.06	0.65
2751	Manufacture of electric domestic appliances	0.04	0.00	0.91	0.11	0.00	0.04	0.48
2511	Manufacture of metal structures and parts of structures	0.03	0.03	0.19	0.03	-0.02	0.00	2.09
2599	Manufacture of other fabricated metal products n.e.c	0.02	0.01	0.29	0.04	0.00	0.00	0.62
2351	Manufacture of cement	0.01	0.00	0.34	0.05	0.01	-0.01	0.24
2910	Manufacture of motor vehicles	0.01	0.00	0.51	0.04	0.01	0.00	0.79
2899	Manufacture of other special-purpose machinery n.e.c	0.002	0.00	0.10	0.01	0.01	0.01	0.20



**Table 3** (continued)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NACE	Label	Mean	Median	Max	Standard deviation	Change 1995–2015	Change 2001–2015	Share of green production
2711	Manufacture of electric motors, generators and transformers	0.0005	0.00	0.04	0.00	0.00	0.00	0.03

Authors' elaboration on PRODCOM data. Production values are deflated to have data at constant prices, with 2010 as base year. The definition of green products used here is explained in Sect. 2 and it is the one called PRODCOM in Fig. 1. Average, median, maximum and standard deviation are computed over all available countries and years (1995–2015), columns 7 and 8 report changes in the average green share for 1995–2015 and 2001–2015 respectively. The last column reports for each industry the share it represents in total green production across all industries, countries and years

industries, the average green share is above 20%, there is always at least one industry with a country-year observation with 100% green production, and the absolute long-term changes are positive, with the exception of production in railway and non-domestic cooling and ventilation. Finally, these industries account for 73.9% of total green production. We then observe a second group of five industries, including the production of LEDs and PV panels (in “electronic components manufacturing”) and wind turbines (manufacturing of engines and turbines), that represent another 21% of the total green production. The remaining eight industries account for just 5.1% of the total green production and always have mean shares of green production below 0.04; thus, we define them as marginally green. In the remainder of the paper, we study green specialisation focusing on the 13 industries identified in Table 3 as high-green-potential industries. However, for the sake of completeness, Online Appendix D replicates the main analyses of Sect. 4 for the full set of green industries (in particular, see Figs. D.1, D.2 and D.3).

## 4 Specialisation Patterns in Green Production

We begin by exploiting the cross-country variation of our data to study specialisation in green production across countries and high-green-potential industries. Our analysis devotes specific attention to the identification of green leaders and the persistency of their comparative advantage. Figure 2 plots the evolution of the 3-years moving average of countries’ green production shares. We group countries based on size and geographic position to look at large (panel A), small (panel B) and Eastern European (panel C) countries, and always including the benchmark of the European (weighted on turnover) average across all available countries in each year.<sup>10</sup>

Green production shares in high-green-potential industries rarely exceed 4% of countries’ total production, with an average just above 2%. This is consistent with the most reliable estimates of the green economy for the US, e.g., Elliott and Lindley 2017; Vona et al. 2019.<sup>11</sup> In terms of country rankings, those with the largest shares of green production are Denmark, Germany, the UK, Sweden and Austria. All leaders are high-income countries that are at the technological frontier and have strong capabilities in high-tech industries.<sup>12</sup> This suggests that engineering and technical competences, which are typically core capabilities for high-green-potential industries, can be reused in green production as we will also show in the econometric analysis.

Not surprisingly, we also find high persistency in the shares of green production that increases by a modest 12.5% over the period between 1995 and 2015 (from 0.02 to 0.0225). Explaining the slow diffusion of green production is beyond the scope of this work, but the rapid rise of China as a manufacturing powerhouse can contribute to explain this pattern

<sup>10</sup> Note that Fig. 2 reports country-level shares, while Table 3 reports the shares of green production within each industry.

<sup>11</sup> As mentioned above, using the broader CLEG green list, we obtain an EU average green share of production around 10% (Figure D.6 in the Online Appendix), which is off target relative to the US benchmark. Note that the measures for the US rely on different methodologies – using either surveys or employment data and also include the service sector. For this reason, direct comparisons between our results and the existing measures of the US are not possible, although the fact that we find similar results brings additional support to our methodology.

<sup>12</sup> We explore in greater detail countries’ green specialisation at the product level in Online Appendix C, Table C1. We find that the top three green products are quite similar across countries.

(Algieri et al. 2011; Sawhney and Kahn 2012; Liu and Goldstein 2013). It is however worth pointing out that such modest increase in shares of green production at the country level masks a quite large growth in absolute terms. On average green industries have seen an increase in sold production of over €136 M over the period 2005–2015. This is largely driven, however, by very high values at the top of the distribution.

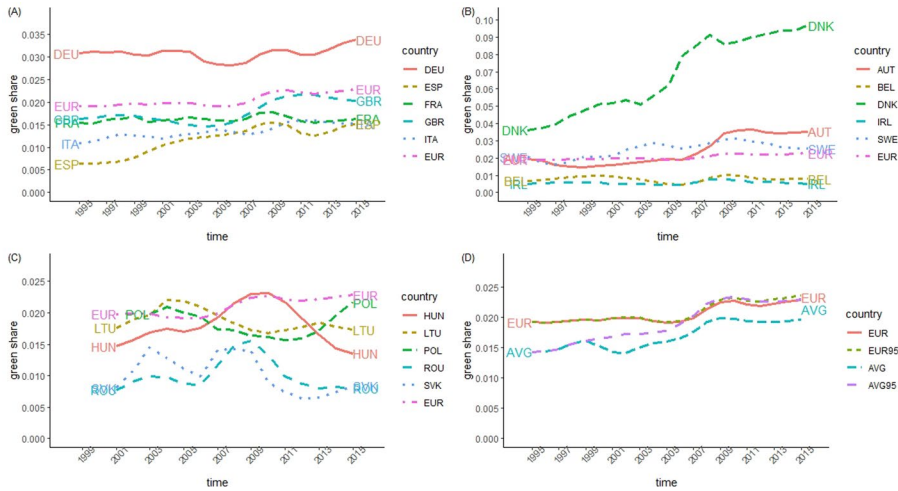
Green production shares are not a direct measure of the extent to which a country specializes into green production since they lack a benchmark for comparison. To this aim, we use a Balassa index of Revealed Comparative Advantage (RCA) index that is widely used to study specialisation patterns (Balassa 1965; Cole et al. 2005; Hidalgo et al. 2007; Petralia et al. 2017). The RCA index is computed as follows:

$$RCA_{it}^{green} = \frac{y_{it}^{green} / y_{it}^{total}}{\sum_i y_{it}^{green} / \sum_i y_{it}^{total}} \quad (2)$$

where  $y_{it}^{green}$  is the green production in country  $i$ . The index compares the green production share of country  $i$  with the average green production share across all countries. Note that RCA values between 0 and 1 denote non-specialisation, while RCA values above 1 denote specialisation. As a result of this asymmetry, statistical analyses using the Balassa's RCA index tends to assign too much weight to values above one (Dalum et al. 1998; Cole et al. 2005; Yu et al. 2009). Laursen (1998) proposes to either take the logarithm of the RCA or to make it varying symmetrically between  $-1$  and  $+1$  applying the following transformation:  $SRCA_{it}^{green} = (RCA_{it}^{green} - 1) / (RCA_{it}^{green} + 1)$ . We use the symmetrical RCA for descriptive purposes as it allows a better visualization of the results, while we resort to the logarithm transformation in our econometric analysis to interpret the results in terms of elasticities. To begin exploring the driver of green specialisation, we correlate countries' green RCA<sup>13</sup> and the OECD index of environmental policy stringency (EPS henceforth) for market-based policies. In Fig. 3, we plot these correlations for selected years, the take-away is that the unconditional correlation between green RCA and EPS is strong and positive, but slightly decreasing over time.<sup>14</sup> In 2001, the EPS index exhibits a correlation of 0.34 with the green RCA, which decreased to 0.19 in 2015. The fact that the strength of this relationship decreases over time will be further explored in the econometric analysis in the next section. Note that we choose market-based policy because economic theory argues in favour of their higher effectiveness in stimulating the diffusion of green goods and technologies (Requate 2005; Nordhaus 2019). Although the empirical literature offers no clear evidence on the superiority of market-based policies over non-market-based ones, at least on innovation outcomes (e.g., Nesta et al. 2018), there are also practical reasons for our choice to use market-based policies. In fact, the non-market-based EPS index shows very

<sup>13</sup> The green RCA is computed by treating green production from high-green-potential industries as a unique sector, i.e.,  $y_{it}$  is the total green production from all high-green-potential industries for each country  $i$ .

<sup>14</sup> In Online Appendix D, we also show the same figure using CLEG products to identify green production. We find that the relationship between green specialisation and environmental policy is weaker, with correlation coefficients between 0.09 and 0.14. This lower sensitivity to environmental policies further supports the exclusion of multiple-usage products.



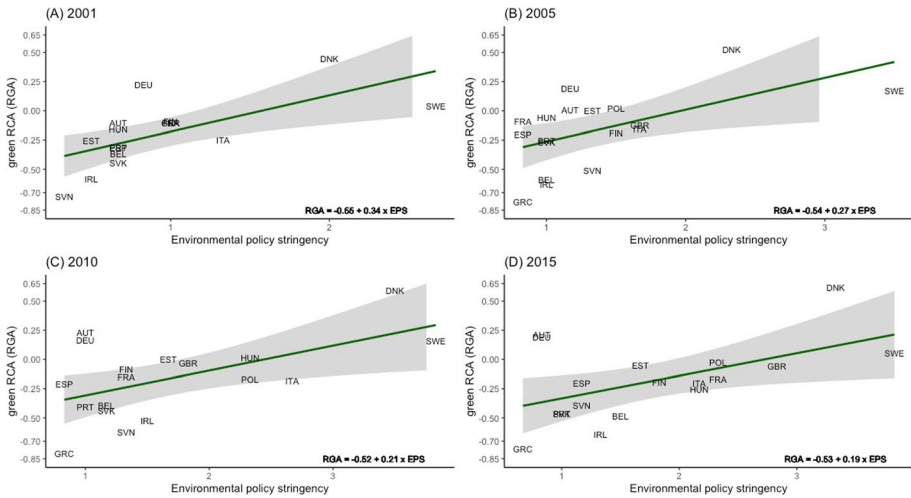
**Fig. 2** Evolution of green production shares for selected European countries. *Notes* Panel A, B and C report green production shares over time for large, small and Eastern European countries, respectively. These have been smoothed by taking 3-years moving averages. Production values are deflated to have data at constant prices, with 2010 as base year. We only use green production from high-green potential industries as identified in Table 3. EUR is the European green shares across all available countries in each year. In panel D, we compare it with the unweighted average (AVG) across countries. Because data on Eastern countries is available only from 2001 onwards, and 2003 onwards for Poland, we report both these measures computed for each year for all available countries as well as only for countries for which we have a balanced panel since 1995, i.e.: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Portugal, Spain, Sweden, United Kingdom, (EUR95 and AVG95, respectively)

little variation across EU countries, making it difficult to use it to identify a policy effect in a multivariate regression.<sup>15</sup>

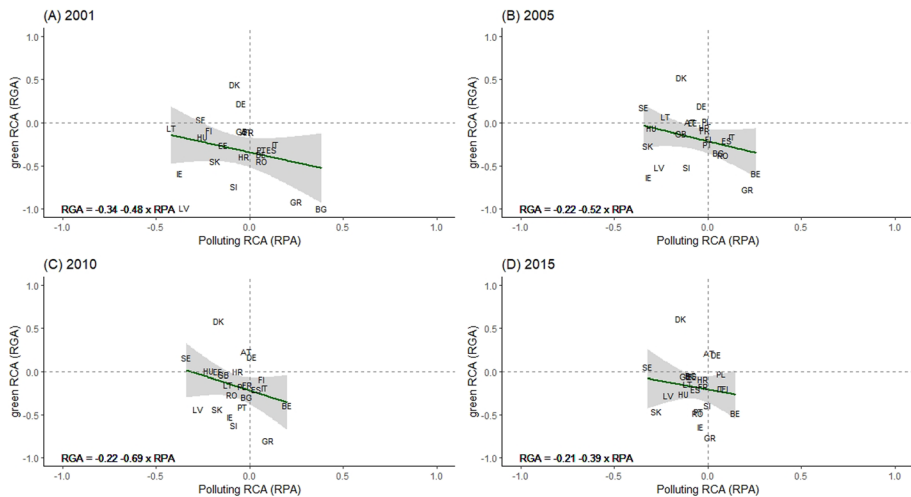
Finally, we also correlate the green RCA with the brown RCA to assess the extent to which there is an overlapping between green and brown specialisation. The brown RCA is computed by treating all polluting industries defined in Table 2 as a single sector and by considering all production as “polluting”. In Fig. 4, we plot green and brown RCA for selected years dividing countries into four quadrants. We choose 2001 as our earliest year because the PRODCOM data are not available for Eastern European countries in previous years. Countries in the top-left quadrant have an RCA in green production but not in polluting production. The top-right quadrant shows countries with an RCA in both types of production, the bottom-right shows countries with an RCA only in polluting production and the bottom-left shows countries with an RCA in neither type of production.

The number of countries with a green RCA (i.e., those above the horizontal dashed line) remains quite stable (with Austria joining Sweden, Germany and Denmark), with only Denmark experiencing a noticeable increase over time. Specialisation in polluting industries shows less dispersion than green specialisation, with most countries clustered around 0 (the vertical dashed line). Brown specialisation emerges in countries with lower income per capita (such as Romania, Bulgaria, Greece) as well as in some traditionally industrial economies (such as Italy and Belgium). Consequently, the green and brown RCAs exhibits

<sup>15</sup> To give an example, in 2015 all countries in our sample have non-market EPS is equal to 5.5 except for Portugal and Italy (which have values of 5 and 6 respectively).



**Fig. 3** Green RCA and environmental policy stringency across countries and over time. *Notes* Authors' elaboration on PRODCOM data and OECD for the index of environmental policy stringency (EPS) for market-based policies. We plot countries' green RCA and the EPS, developed by the OECD. Green RCA is based solely on green production from high-green potential industries, as identified in Table 3. Production values are deflated to have data at constant prices, with 2010 as base year. The RCAs are computed following Eq. 2 are made symmetrical around 0 and bounded between -1 and 1, the value of 0 indicates therefore whether a country has successfully specialised in green production. We also report the coefficient of a regression of green RCA on the EPS index for each year



**Fig. 4** Green and polluting RCA across countries and over time. *Notes* Authors' elaboration on PRODCOM data. We plot countries' green and polluting RCA. Green RCA are based solely on green production from high-green potential industries, as identified in Table 3. Polluting production is total production from polluting industries identified in Table 2. Production values are deflated to have data at constant prices, with 2010 as base year. The RCAs are computed following Eq. 2 are made symmetrical around 0 and bounded between -1 and 1, the value of 0 indicates therefore whether a country has successfully specialised in green production. We also report the coefficient of a regression of green RCA on polluting RCA for each year

a strong negative correlation, i.e., the correlation coefficient is always beyond  $-0.39$ .<sup>16</sup> This evidence, together with the fact that the green leaders are mostly high-income countries, is suggestive on the possible cross-country distributional effects of EU environmental policies, such as the European Green Deal, may exacerbate the gap between the core and the periphery of Europe. It is therefore important, when it comes to providing a large fiscal push for the green economy, that more attention is given to helping laggard countries to develop a comparative advantage in some specific green products. Our econometric analysis of next Section provides further insights on the design of green industrial policies for both laggard and leading countries.

## 5 Drivers of Green Specialisation

To examine econometrically the drivers of green specialisation, our starting point is the canonical empirical framework in the literature on the drivers of specialisation (Midelfart-Knarvik et al. 2000; Romalis 2004; Mulatu et al. 2010; Nicoletti et al. 2020). In its simplest form, this framework compares the influence of two main sources of comparative advantages: (i) abundance of productive factors, stemming from the Heckscher-Ohlin (HO) theoretical framework; (ii) market access and economies of scale, in line with new economic geography (NEG) theory.

The empirical implementation relies on a shift-share measurement framework following the seminal paper of Rajan and Zingales (1998). More specifically, HO- (and NEG-) drivers are included in the analysis by interacting a measure of industry-level intensity of a given productive factor and a measure of country-level abundance of such factor. Endogeneity concerns are mitigated in this framework because industry-level characteristics are taken as cross-country averages and fixed over time, while country-level drivers are allowed to vary over time. However, as well-known in the econometric literature (Angirst and Pischke 2009), solving multiple endogeneity problems is exceedingly difficult in reduced-form regressions because it is not possible to establish a well-defined counterfactual. Our goal here is not to isolate a specific causal channel, but to ascertain the importance of different drivers—which we derive from the theoretical literature—of green specialisation. Therefore, while the trade literature usually gives a causal, theory-driven interpretation to the coefficients estimated through this approach, the estimates presented in this section should be interpreted as theory-driven correlations rather than causal effects.

### 5.1 Empirical Specification

Mulatu et al. (2010) expand the empirical framework of Midelfart-Knarvik et al. (2000) to study the patterns of specialisation of polluting industries and the possible emergence of pollution havens within the EU area. Their key variable of interest is the sectoral pollution (or energy) intensity interacted with proxies of environmental regulation (or energy prices) at the country level. With this aim, the authors consider a cross-section of 16

---

<sup>16</sup> As we show in Online Appendix D, this relationship becomes essentially flat when we use the CLEG list for green products, in line with the fact that with this list there is more overlap between green and polluting productions.

manufacturing industries in 13 countries and exploit across industries variation in pollution exposure to estimate the role of environmental policies on industry location.

Taking stock from the descriptive evidence, we adapt the canonical model to account for the peculiar characteristics of green industrial sectors.

First, green production is highly concentrated in a few high-green potential sectors. In order to avoid misleading comparisons with sectors not involved in the green transition, we limit the analysis of the drivers to high-green potential sectors. Second, we show that the sectors with a high green potential are not among the carbon-intensive ones. Accordingly, we measure exposure to environmental policies based on the sectoral share of green production at the 4-digit industry level rather on the pollution intensity. Finally, because high-green potential sectors are often high-to-medium tech, it is important to take stock from the existing literature on the drivers of specialisation in high-technology green sectors. More specifically, we add three drivers that have been examined by the literature on green innovation using patents: i. path dependency (e.g., Aghion et al. 2016), which is also in line with our descriptive evidence; ii. complementarity with proximate capabilities (e.g., Peruchas et al. 2020), which resonates with the literature showing that specialisation in one product is related to specialisation in other “similar” products (Hidalgo et al. 2007; Mealy and Teytelboym 2020); iii. diversification of the knowledge base (e.g., Colombelli and Quatraro 2019), which increases the scope for recombinant innovations (Weitzman 1998).

To fix the ideas, we begin with an econometric model that mimics the specification of Mulatu et al. (2010) over the period 2005–2015<sup>17</sup>:

$$\ln(RCA_{ij,t}^g) = \alpha + \sum_k \gamma_k \times \bar{\phi}_j \times X_{it} + \tau_t + \varepsilon_{ijt} \quad (3)$$

where  $\varepsilon_{ijt}$  is the error term,  $X_{it}$  are the  $k$  country-level drivers explained below (with  $\gamma_k$  being the associated coefficients), and time dummies  $\tau_t$  absorb common shocks for all EU countries.

Our dependent variable is the index of revealed green comparative advantage index,  $RCA_{ij,t}^g$ . Note that, while in the previous section we used an RCA at the country level, the variable  $RCA_{ij,t}^g$  varies by country  $i$ , sector  $j$  and time  $t$ . In other words, RCA in previous sections considered green production as a single industry. Here we fully exploit our data looking at specialisation in both green and non-green production across countries *and* industries. We compute an RCA at the country-industry level for both green and non-green production, as follows:

$$RCA_{ijt}^k = \frac{y_{ijt}^k / \sum_j y_{ijt}}{\sum_i y_{ijt}^k / \sum_j \sum_i y_{ijt}} \quad (4)$$

<sup>17</sup> Most Eastern European countries enter in our dataset in 2001, except for Poland, which is included only from 2003 onwards. Therefore, focusing on the years 2005–2015 allows us to have a balanced panel and to compute pre-sample means for all countries.



where  $k = g$  (green) or  $ng$  (non-green, i.e.,  $y_{ijt}^{ng} = y_{ijt} - y_{ijt}^g$ ).<sup>18</sup> We refer to these two measures as green and non-green RCA, respectively. Hence, we normalise the share that green production represents in total production in industry  $j$  in country  $i$ , by the share of green production in total production in industry  $j$  across all countries. In our econometric application, we use the log of the asymmetric  $RCA_{ij,t}^g$ : this has the benefit of dealing with the skewedness of the index, reducing its asymmetry (Dalum et al. 1998; Soete and Verspagen 1994).

In this basic specification that closely follows Mulatu et al. (2010), our main explanatory variables combine time-invariant industry characteristics  $\bar{\phi}_j$  and country characteristics  $X_{it}$  that vary over time. These are reported in Table E.3 in the Online Appendix, where we discuss them in detail. In a nutshell we include interaction measures on (i) environmental policies, (ii) capital intensity, (iii) skills, (iv) technology, (v) economies of scale, and (vi) market access, we briefly discuss these in turn.

Concerning environmental policies, we interact the average green production share of the sector in Europe with the OECD EPS index of stringency in market-based policies. Capital intensity is measured as the ratio between investments in tangible assets and total employment of the 4-digit sector (source Structure of Business Survey, SBS, of Eurostat) and it is interacted with the log of the investment in tangible, non-residential, assets over total employment of each country-year (source EUKLEMS-INTANProd data). For sectoral high-skill intensity, we compute the ratio between employment in abstract occupations and total sectoral employment in the US and interact this with the share of workers with tertiary education of each country-year. We account for the technological complexity of production by interacting the share of R&D personnel and researchers in total active population from Eurostat and a dummy taking value 1 for high- and medium-high tech manufacturing industries, following Eurostat's definition based on R&D expenditure, which we report in Table E.5 in the Online Appendix.<sup>19</sup> Finally, we also include proxies for economies of scale and market potential as possible drivers in line with the NEG literature. First, we interact total manufacturing production for each country with the average number of employees per plant of the sector.<sup>20</sup> Second, we use total value added of each country as a measure of market size and industry's share of final goods (either capital or consumption goods) in total production. We refer the reader to Tables E.2 and E.3 in the Online Appendix for detailed descriptive statistics and data sources on these variables.

In our favourite specification, we progressively add to Eq. 3 the three key drivers identified by the literature on green technology as important for green specialisation, i.e., path dependency, complementarity with non-green capabilities and diversification of capabilities. In doing so, we estimate variants of the following equation:

<sup>18</sup> Note that here we only look at high-green potential industries, none of which can be considered as GHG-intensive. Therefore, when we compute green and non-green RCAs at the industry level, we are comparing the green and non-green production *within* the same green industry. By non-green production, we simply refer to the share of production of a given industry of goods that are not part of our favourite list of green products. The non-green RCA is thus different from the polluting RCA of Fig. 5, which is based on the production of GHG-intensive industries shown in Table 2 and computed at the country-level.

<sup>19</sup> We have also replicated this analysis using patents as a share of output, results are robust and available upon request.

<sup>20</sup> Following Mulatu et al. (2010) we argue that the optimal scale of a sector can be inferred using the average number of employees of all firms in that sector across EU countries and interact this variable with manufacturing output to capture the size of production of the country.

$$\ln(RCA_{ij,t}^g) = \alpha + \sum_k \gamma_k \times \bar{\phi}_j \times X_{it} + \sum_t \beta_t \times \ln(\overline{RCA}_{ij,t_0}^g) + \gamma \ln(RCA_{ij,t-1}^{ng}) + \delta \ln(\#RCA_{ij,t-1}^g) + \vartheta \ln(\#RCA_{ij,t-1}^{ng}) + \tau_t + \varepsilon_{ijt} \quad (5)$$

The main proxy of path dependency in green specialisation is the pre-sample mean of the green RCA ( $\overline{RCA}_{ij,t_0}^g$ ) computed for the years 2001–2004. We interact the pre-sample mean of green RCA with time dummies to assess the persistence of the “first-mover advantage”. Using the pre-sample mean of green production is also more consistent with the notion of path dependency than using the lagged dependent variable.

Inspired by the recent literature that uses the product space to map products’ similarities (Hidalgo et al. 2007; Mealy and Teytelboym 2020), we measure the degree of complementarity between green and non-green capabilities using the level of non-green RCA within the same four-digit sector and lagged by 1 year ( $RCA_{ij,t-1}^{ng}$ ).<sup>21</sup> Taking the level of non-green specialisation within the same detailed 4-digit sector represents a natural way to measure capabilities that are similar to green ones. The effect of having a stronger non-green RCA on green specialisation is unclear ex-ante. It could be positive if the non-green capabilities can be replicated and successfully used to create a green comparative advantage within the same sector. It could be negative if there is competition between the green and non-green uses of a similar pool of capabilities. While determining which effect would prevail is an empirical issue that we will explore through Eq. 5, the unconditional correlation between green  $RCA_{ij,t}^g$  and non-green  $RCA_{ij,t}^{ng}$  is rather high (0.5). Thus, we expect stronger non-green capabilities within the same sector to be positively associated with the building of a comparative advantage in green industries.

Finally, we capture green (non-green) diversification in a country’s capabilities within a particular sector with the number of green (non-green) products with an  $RCA > 1$ , i.e., above the threshold designating a country as having a comparative advantage for that product at time  $t - 1$  ( $\#RCA_{ij,t-1}^g$  and  $\#RCA_{ij,t-1}^{ng}$  for green and non-green diversification, respectively). To account for skewness in this measure, we take the log of both variables.<sup>22</sup> We argue, in line with the well-established literature on structural change, that countries specialise in products based on their productive capabilities (Hidalgo et al. 2007; Hidalgo and Hausmann 2009), and therefore, the number of green goods produced with an RCA within each country-industry, will capture the breadth of green productive capabilities.

## 5.2 Results

Table 4 contains the main result of our econometric analysis. We begin with the specification of Eq. 3, then we progressively add the other drivers included in Eq. 5. The last column (5) is our favourite specification where we also include country fixed effects to Eq. 5 in order to account for unobservable differences in policies and institutions that may be

<sup>21</sup> We use this approach rather than building a fully-fledged measure of product proximity based on product-space approaches (see Hidalgo et al. 2007). This is because such proximity measures are built using co-occurrence in green and non-green RCA. This makes the approach not suitable to be used in an econometric analysis of the drivers at the sector-by-country level, since correlation between green and non-green specialisation would be established by construction.

<sup>22</sup> We deal with the case in which the number of products is 0 by adding 1 so that the log transformation does not yield missing values.

**Table 4** Drivers of green specialisation

	(1)	(2)	(3)	(4)	(5)
Country environmental policies	0.161 (0.450)	-0.483 (0.322)	-0.351 (0.320)	-0.367 (0.276)	-0.0929 (0.196)
Industry share of green production	-3.067 (2.256)	-2.168 (1.388)	-0.331 (1.386)	-0.814 (1.228)	-1.405 (1.197)
Country env. Policies * Industry share of green production	1.644* (0.984)	1.190* (0.706)	0.656 (0.675)	0.555 (0.558)	0.672 (0.555)
Population (log)	4.444*** (0.950)	2.015*** (0.715)	1.715** (0.668)	1.679*** (0.604)	3.030 (4.475)
Country capital intensity	0.481 (1.849)	1.823 (1.164)	1.694 (1.159)	1.453 (1.064)	1.157 (1.368)
Industry capital intensity	-0.914 (0.634)	0.227 (0.452)	0.291 (0.432)	0.214 (0.427)	0.102 (0.412)
Country capital intensity * Industry capital intensity	0.214 (0.334)	-0.192 (0.233)	-0.207 (0.227)	-0.157 (0.219)	-0.125 (0.210)
Country skills	-0.459 (1.478)	-0.877 (0.952)	-0.819 (0.921)	-0.669 (0.840)	-0.751 (1.016)
Industry skills	-0.149 (0.300)	-0.0815 (0.165)	-0.120 (0.167)	-0.166 (0.155)	-0.146 (0.142)
Country skills * Industry skills	0.0256 (0.0882)	0.0246 (0.0508)	0.0355 (0.0507)	0.0403 (0.0469)	0.0340 (0.0433)
Country technology	1.493 (1.310)	0.594 (0.861)	0.472 (0.814)	0.562 (0.746)	0.199 (0.738)
Industry technology	6.175 (5.252)	2.316 (3.174)	2.048 (2.889)	0.324 (2.653)	0.864 (2.481)
Country technology * Industry technology	1.230 (1.125)	0.532 (0.697)	0.493 (0.645)	0.0869 (0.587)	0.194 (0.546)
Country scale	-2.176** (1.103)	-1.176* (0.616)	-0.967* (0.573)	-0.868 (0.535)	-0.581 (1.162)
Industry scale	-10.85 (6.849)	0.129 (3.552)	0.806 (3.438)	2.181 (3.041)	1.898 (2.436)
Country scale * Industry scale	0.386 (0.258)	-0.0214 (0.135)	-0.0476 (0.130)	-0.0983 (0.115)	-0.0899 (0.0924)
Country market access	-2.479*** (0.736)	-0.411 (0.443)	-0.330 (0.429)	-0.199 (0.406)	0.853 (1.580)
Industry market access	-1.761 (11.44)	9.025 (7.512)	8.234 (7.151)	13.36* (6.981)	12.67** (6.204)
Country market access * Industry market access	0.0385 (0.424)	-0.344 (0.274)	-0.309 (0.261)	-0.534** (0.259)	-0.509** (0.230)
Non-green RCA (log, t-1)		0.917*** (0.0372)	0.831*** (0.0509)	0.733*** (0.0532)	0.729*** (0.0536)
Number of green products with RCA (log, t-1)		0.872*** (0.0423)	0.789*** (0.058)	0.697*** (0.061)	0.692*** (0.0605)
Number of non-green products with RCA (log t-1)		0.796*** (0.0569)	0.707*** (0.0736)	0.622*** (0.0762)	0.619*** (0.0787)
2005 * PSM RGA (ln)		0.792*** (0.0597)	0.714*** (0.0734)	0.629*** (0.0747)	0.625*** (0.0778)
2006 * PSM RGA (ln)		0.759*** (0.0598)	0.687*** (0.0717)	0.602*** (0.0702)	0.593*** (0.073)
2007 * PSM RGA (ln)		0.764*** (0.0593)	0.706*** (0.0654)	0.613*** (0.067)	0.597*** (0.07)
2008 * PSM RGA (ln)					
2009 * PSM RGA (ln)					
2010 * PSM RGA (ln)					

Table 4 (continued)

	(1)	(2)	(3)	(4)	(5)
2011 * PSM RGA (ln)		0.779*** (0.059)	0.728*** (0.063)	0.635*** (0.0648)	0.613*** (0.0689)
2012 * PSM RGA (ln)		0.755*** -(0.063)	0.702*** (0.0661)	0.618*** (0.0649)	0.596*** (0.0683)
2013 * PSM RGA (ln)		0.720*** (0.0663)	0.659*** (0.071)	0.576*** (0.0719)	0.550*** (0.0747)
2014 * PSM RGA (ln)		0.687*** (0.0714)	0.627*** (0.0737)	0.548*** (0.0756)	0.517*** (0.0772)
2015 * PSM RGA (ln)		0.681*** (0.0699)	0.616*** (0.0738)	0.541*** (0.0738)	0.516*** (0.0737)
Constant	60.01** (29.10)	12.75 (19.24)	9.128 (18.36)	3.534 (16.61)	-54.13 (93.22)
<i>Observations</i>					
R-squared	2444	2444	2444	2444	2444
Year FE	0.256	0.637	0.651	0.694	0.714
Country FE	Yes	Yes	Yes	Yes	Yes

Pre-sample mean computed for the years 2001–2004, for Poland only for 2003–2004 due to data constraints. Production values are deflated to have data at constant prices, with 2010 as base year. All RGA explanatory variables, except the pre-sample mean, are lagged by 1 year. Variable names are simplified for space's sake, Table E.3 reports full details. Estimation time span is 2005–2015. Robust standard errors in parentheses \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

correlated with the green RCA. As in Mulatu et al. (2010), for the variables of the canonical model the interpretation requires computing the marginal effect of country-level driver (e.g., university graduates) in correspondence to different percentiles of industry-level characteristics (i.e. share of highly skilled workers). For sake of space, we present such calculations in Table E.12 and E.14 of the Online Appendix, for both our favourite specification of column 5 and the canonical specification of column 1.

Column 1 presents the results of the canonical model. The bottom line is that none of the standard drivers matter for green specialisation. This conclusion is confirmed when we compute the marginal effects of the drivers at different percentiles of industry characteristics (Table E.13). Two exceptions are present. First, the EPS index, for which we find that the marginal effect increases together with the share of green production of the industry, suggesting that green policies may at best reinforce existing patterns of green specialisation. While the effect of the interaction term is statistically significant at the 10% level only in the basic specification of Column 1, Table E.12 reporting marginal effects shows that the EPS index becomes statistically significant already at the median of the green industry share. However, the association between the EPS and the green RCA becomes smaller and statistically insignificant when adding the controls for path-dependency, non-green capabilities, and diversification (columns 2–5). In our favourite specification from column 5, Table E.13 shows that the EPS never passes the threshold above which its effect on green specialisation becomes statistically significant, thus the effect of environmental policies seems to pass through that of other structural factors that predate the time span considered in our analysis.

To further explore the role of environmental policies, Table E.6 in the Online Appendix E presents the correlation between the pre-sample mean of green specialisation and the EPS index, which is quite large (0.267) and statistically significant at the conventional level. We also note that the market-based EPS index is correlated more strongly with the pre-sample mean green specialisation than the non-market-based index (0.097). This implies that environmental policies are particularly effective on green specialisation when they are used to build an early start advantage and that the market-based EPS is more correlated with such advantage in our sample.

This result is also evident in Table E.11 of the Online Appendix E where we estimate Eqs. 3 and 5 separately for the 3 years (2005, 2010, 2015), thus only exploiting the cross-sectional variation of the data as in Mulatu et al. (2010). The interaction term between the EPS index and the share of green production is statistically significant only in the first period (2005). Interestingly, the interaction term remains positive and statistically significant in 2005 also for the specification of Eq. 5. These robustness checks allow to qualify our results concerning the effect of environmental policies on green specialisation by showing the higher effectiveness of such policies to build an early advantage in green production.

Concerning the other drivers, we find that, somewhat surprisingly, the interaction for market access shows a negative coefficient that does not vary with the industry share of final goods in total output in Table E.12. This is difficult to interpret because of high collinearity with the other variables used to proxy for scale effects. However, it should be noted that when we move to our favourite specification, in Table E.13, the marginal effects become statistically insignificant suggesting that other structural factors fully absorb possible effects of market access. Two, not mutually exclusive but untestable, explanations can account for this result. First, there is measurement error in our proxies of the drivers that leads to an attenuation bias. The skill intensity, for instance, is obtained from US data through a crosswalk between the US and the EU industry classification that includes

several many-to-many matches and is not perfect. Second, the variation in the industry characteristics is much smaller in our subsample of high-green potential industries compared to the sample of all manufacturing industries used by Mulatu et al. (2010). Table E.5 confirms this conjecture by reporting the standard deviation and the coefficient of variation of three industry characteristics (skill, capital intensity and the average number of employees per plant) for high-green potential industries and all manufacturing industries. The subsample of high-green potential industries exhibits less variability in the industry characteristics (which are used to estimate the effect of country-level drivers) compared to the entire sample of manufacturing industries.

Column 2 considers a specification where we add the pre-sample mean of the green RCA, which is our proxy for first-mover advantage. The pre-sample mean is interacted with time dummies to estimate the speed at which the pre-2004 advantage fades away. The persistency of such advantage is remarkable: the elasticity of the pre-2004 green RCA is 0.92 after 1 year and 0.68 after 11 years. As we progressively add other variables in columns 3, 4 and 5, we observe a concomitant decline in the influence of the first-mover advantage. In the most comprehensive specification of column 5, the elasticity of the initial advantage is still 0.73 after 1 year and 0.52 after 11 years. This implies that, conditional on the other covariates that are also correlated with long-term structural factors and thus with first-mover advantage, a one standard deviation in the log of the initial green advantage (3.31, see Table E.1) continues to explain as much as a 50.9% of one standard deviation in the log of the green RCA (3.39, see Table E.1) after 11 years.<sup>23</sup>

Overall, the first policy implication of our analysis is that green specialisation is highly persistent and policy actions play at best a minor role in reverting pre-existing patterns of specialisation in green industries.

In column 3, we highlight our second main finding: green and non-green specialisation reinforce each other as highlighted by the positive and statistically significant coefficient of the non-green RCA. The quantitative impact is also non-negligible: a one standard deviation change in the log of non-green specialisation (2.855) explains as much as 15.7% of one standard deviation in the log of green RCA (3.388). This association becomes quantitatively smaller by progressively adding proxies of diversification (column 4) and country fixed effect (column 5). In this most comprehensive specification of column 5, the non-green RCA still explains 9.1% of one standard deviation of the green RCA. The similarity between green and non-green competences within a narrowly defined domain resonates with previous findings of Perruchas et al. (2020) using patents, Mealy and Teytelboym (2020) using export data and Vona et al. (2018) using skill data. Because building such capabilities takes time and depends on structural characteristics of a country's industrial system, the complementarity of green and non-green capabilities helps explain the persistency in green specialisation.

In column 4, we include the proxies of diversification in green and non-green productions. We find that the number of green products with an RCA is positively and significantly correlated with the average green RCA of the industry. A one standard deviation change in the number of green products with RCA explains as much as 23.7% of a standard deviation in the green RCA. Note that the number of green products with RCA is correlated with the lagged green RCA, thus it mechanically captures part of the path-dependency

<sup>23</sup> This number is obtained by multiplying the coefficient of the log of the green RCA PSM after 11 years (0.52) by its standard deviation (3.31) and dividing it by a standard deviation in the log of the green RCA (3.39).

effect. Furthermore, the coefficient of non-green diversification is far from being statistically significant at conventional level. Since the number of non-green products with an RCA is mechanically correlated with the non-green RCA, it is not surprising to detect a decline in the coefficient associated with the non-green RCA in columns 4 and 5. This last set of results suggests that diversifying the set of capabilities is important for maintaining a comparative advantage in green production. Yet, because only a few countries have multiple green products with a revealed comparative advantage in a specific sector, the diversification channel is not easily accessible for laggard countries that want to catch up with leaders.

In the Online Appendix, we conduct a series of robustness checks of these results. First, results hold when we consider all green sectors (Table E.7). The main notable difference is the positive and significant effect of the interaction between environmental policies and the average green share of production of the industry in the favourite specification of column 5. Thus, a policy stimulus becomes more effective if there is more variability in the set of industries included in the estimation sample. This also suggests that, while structural determinants are particularly relevant for high-green potential industries, policy drivers gain importance for marginally green industries. Second, weighting the regressions using the average industry turnover does not alter the main results, but again reinforces the effect of the EPS index in greener sectors (Table E8). The effect of the non-green RCA is estimated less precisely in our favourite specification, leading to non-statistically significant relationship between green and non-green specialisation. In addition, we conduct the same analysis using the CLEG list (Table E.9). Importantly, results on the main drivers (pre-sample mean of the green RCA, non-green RCA and diversification proxies) are qualitatively similar to those obtained using our favourite list, although the estimated elasticities are somewhat smaller. Finally, we explore an alternative functional specification for the number of green (and non-green) products with RCA to account for the high skewedness of this variable. We replace the continuous variable with three dummies: no product with RCA, one product with RCA and at least two products with RCA (Table E.10). Including these variables do not alter the main results, but for the positive correlation between green and non-green specialisation that becomes statistically insignificant at conventional level. In turn, the dummies associated with non-green diversification becomes statistically significant. Given the strong collinearity between these two variables, we interpret these results as an indication of the importance of non-green capabilities for green specialisation without taking a strong position on channel through which this influence takes place.

## 6 Conclusions

This paper presents new stylized facts on the structure and evolution of specialisation in green productions by assembling a new dataset based on the PRODCOM dataset of Eurostat, which allows for the first time to examine variation in green production across detailed sectors (4-digit NACE), countries (in the EU) and over several years (1995–2015). We construct a favourite list of green products by cleaning existing lists proposed during recent international negotiations to reduce trade tariffs for such products. Our main criterion is to exclude green goods with double usages from our final list because this is the most controversial issue in the debate on the definition of what should be considered as green.

Our first finding is that there is no overlap between green production and the (direct and indirect) GHG-intensity across two-digit NACE industries. This result has two important



implications. In the debate on the definition of what is green, the process- and output-based approaches capture different aspects of the green economy. Naturally, both definitions are important to understand the green transition. The paper strives to examine specialisation patterns using both definitions, although data constraints on emissions content of production only make this possible at a coarser level of aggregation (2 digits of NACE) than what is available within PRODCOM. In spite of this limitation, the analysis of the revealed green and brown comparative advantages indicate that European countries tend to specialise either in green or brown sectors.

Exploiting the granularity of our data at 4-digit industry-level, the second result is that green production is highly concentrated in a few sectors despite an average increase of 12.5% (from 2 to 2.25%) over the considered period: out of 119 4-digit manufacturing sectors, 13 of them represent 95% of green production among EU countries.

Third, we rely on revealed comparative advantage measures and find that that green leaders are high-income countries where high-to-medium tech manufacturing industries are traditionally strong and environmental policies more stringent. Taken together with the divergent country specialisation on green and brown sectors, this result raises the concern that the EU green deal plan may exacerbate existing cross-country inequalities.

Last, we examine the drivers of green specialisation comparing the role of standard drivers considered by the literatures on trade specialisation and on environmental innovation. Our results highlight a remarkable persistence in green specialisation suggesting that first-mover advantage is an important factor at play. Moreover, within similar 4-digit industries, green and non-green specialisations complement and reinforce each other. The role of such complementarities is clearly smaller than that of path dependency but corroborates the descriptive analysis pointing to the pre-existing advantage in certain high-to-medium tech sectors. Diversifying the portfolio of green products with comparative advantage is important for sustaining green specialisation. Finally, our analysis suggests that environmental policies are more effective to building a lead start advantage in green industries rather than to creating a new comparative advantage for laggard countries.

A shortcoming of our analysis is that the data are limited to European countries. Because the index of comparative advantage is relative in nature and depends on the number of countries available in the data, there is limited cross-country variation in our data. This is compensated by the fact that we can study production at a highly detailed level of resolution and that our data include all production and not just export flows.

Another limitation of the PRODCOM data is that it only covers the production of manufactured goods and thus excludes the service sector. Leaving services out of our analysis means ignoring the largest part of European economies, some of which, such as knowledge intensive business sectors, may have a significant enabling role in the green economy. Finally, our analysis identifies green products based on their potential to benefit the environment, and comparison with pollution intensity production is possible only at 2-digit level of aggregation. Future research will greatly benefit from more disaggregated information on the pollution content of production so that both output and process approaches can be used within the same analytical framework.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10640-023-00781-7>.

**Acknowledgements** This work was supported by Horizon 2020 Framework Programme, project INNO-PATHS [Grant Number 730403]. We thank the editor, Ian Bateman, and three anonymous reviewers for their extremely constructive and insightful comments. David Popp for useful suggestions and Tobias Kruse

for kindly providing us the updated data on environmental policy stringency of the OECD. Usual disclaimer applies.

**Funding** Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aghion P, Dechezleprêtre A, Hemous D, Martin R, Van Reenen J (2016) Carbon taxes, path dependency, and directed technical change: evidence from the auto industry. *J Polit Econ* 124(1):1–51
- Agrawala S, Dussaux D, Monti N (2020) What policies for greening the crisis response and economic recovery? Lessons learned from past green stimulus measures and implications for the COVID-19 crisis. OECD Environment Working Papers No. 164
- Algieri B, Aquino A, Succurro M (2011) Going “green”: trade specialisation dynamics in the solar photovoltaic sector. *Energy Policy* 39(11):7275–7283
- Ambec S, Cohen MA, Elgie S, Lanoie P (2013) The Porter hypothesis at 20: can environmental regulation enhance innovation and competitiveness? *Rev Environ Econ Policy* 7(1):2–22
- Angrist JD, Pischke JS (2009) Instrumental variables in action: sometimes you get what you need. Mostly harmless econometrics: an empiricist's companion pp. 113–220.
- APEC (2012) APEC Leaders' Declaration ANNEX C APEC List of Environmental Goods, Vladivostok, 8–9 September
- Arrow KJ (1971) The economic implications of learning by doing. In: Hahn FH (ed) *Readings in the theory of growth*. Palgrave Macmillan, London, pp 131–149
- Balassa B (1965) Trade Liberalization and Revealed Comparative Advantage. *Manch Sch Econ Soc Stud* 33:99–123
- Balconi M, Pozzali A, Viale R (2007) The “codification debate” revisited: a conceptual framework to analyze the role of tacit knowledge in economics. *Ind Corp Chang* 16(5):823–849
- Balineau G, de Melo J (2011) Stalemate at the negotiations on environmental goods and services at the Doha Round. *Fondation Pour Les Études Et Recherches sur le Développement International Working Paper No. 28*
- Barbieri N, Perruchas F, Consoli D (2020) Specialisation, diversification, and environmental technology life cycle. *Econ Geogr* 96(2):161–186
- Barrows G, Ollivier H (2018) Cleaner firms or cleaner products? How product mix shapes emission intensity from manufacturing. *J Environ Econ Manag* 88:134–158
- Bernard AB, Jensen JB, Redding SJ, Schott PK (2007) Firms in international trade. *J Econ Persp* 21(3):105–130
- Bernard AB, Jensen JB, Redding SJ, Schott PK (2012) The empirics of firm heterogeneity and international trade. *Annu Rev Econ* 4(1):283–313
- Bombardini M, Gallipoli G, Pupato G (2012) Skill dispersion and trade flows. *Am Econ Rev* 102(5):2327–2348
- Bontadini F, Corrado C, Haskel J, Iommi M, Jona-Lasinio C (2023) EUKLEMS & INTANProd: industry productivity accounts with intangibles. In: *Sources of growth and productivity trends: methods and main measurement challenges*, Luiss Lab of European Economics, Rome
- Bovenberg AL, Smulders S (1995) Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model. *J Public Econ* 57(3):369–391
- Calel R, Dechezleprêtre A (2016) Environmental policy and directed technological change: evidence from the European carbon market. *Rev Econ Stat* 98(1):173–191
- Cantore N, Cheng CFC (2018) International trade of environmental goods in gravity models. *J Environ Manag* 223:1047–1060

- Chen Z, Marin G, Popp D, Vona F (2020) Green stimulus in a post-pandemic recovery: the role of skills for a resilient recovery. *Environ Resource Econ* 76(4):901–911
- Clarke AJ (2006) Learning-by-doing and aggregate fluctuations: does the form of the accumulation technology matter? *Econ Lett* 92(3):434–439
- Cole MA, Elliott RJ, Shimamoto K (2005) Why the grass is not always greener: the competing effects of environmental regulations and factor intensities on US specialisation. *Ecol Econ* 54(1):95–109
- Colombelli A, Quatraro F (2019) Green start-ups and local knowledge spillovers from clean and dirty technologies. *Small Bus Econ* 52(4):773–792
- Cowan R, David PA, Foray D (2000) The explicit economics of knowledge codification and tacitness. *Ind Corp Chang* 9(2):211–253
- Dalum B, Laursen K, Villumsen G (1998) Structural change in OECD export specialisation patterns: de-specialisation and 'stickiness.' *Int Rev Appl Econ* 12(3):423–443
- de Melo J, Solleder J (2018) The EGA Negotiations: why they are important, why they are stalled and challenges ahead. *Fondation Pour Les Études Et Recherches sur le Développement International Working Paper No. 236*
- de Vries GJ, Ferrarini B (2017) What accounts for the growth of carbon dioxide emissions in advanced and emerging economies? The role of consumption, technology and global supply chain participation. *Ecol Econ* 132:213–223
- Dechezleprêtre A, Sato M (2017) The impacts of environmental regulations on competitiveness. *Rev Environ Econ Policy* 11(2):183–206
- Dutz MA, Sharma S (2012) Green growth, technology and innovation. *The World Bank*
- Elliott RJ, Lindley JK (2017) Environmental jobs and growth in the United States. *Ecol Econ* 132:232–244
- European Commission (2019) A6 a balanced and progressive trade policy to harness globalisation
- Eurostat (2015) Glossary: High-tech classification of manufacturing industries. *Eurostat Statistics Explained*
- Eurostat (2016) Environmental goods and services sector accounts manual: 2016 edition. Eurostat, Luxembourg
- Fankhauser S, Bowen A, Calel R, Dechezleprêtre A, Grover D, Rydge J, Sato M (2013) Who will win the green race? In search of environmental competitiveness and innovation. *Glob Environ Change* 23(5):902–913
- Fraccascia L, Giannoccaro I, Albino V (2018) Green product development: what does the country product space imply? *J Clean Prod* 170:1076–1088
- Frondel J, Horbach J, Rennings K (2007) End-of-pipe or cleaner production? An empirical comparison of environmental innovation decisions across OECD countries. *Bus Strateg Environ* 16(8):571–584
- Galindo-Rueda F, Verger F (2016) OECD taxonomy of economic activities based on R&D intensity. *OECD Publishing, Paris*
- Hatch NW, Mowery DC (1998) Process innovation and learning by doing in semiconductor manufacturing. *Manag Sci* 44(11-part-1):1461–1477
- He Q, Fang H, Wang M, Peng B (2015) Trade liberalization and trade performance of environmental goods: evidence from Asia-Pacific economic cooperation members. *Appl Econ* 47(29):3021–3039
- Hidalgo CA, Hausmann R (2009) The building blocks of economic complexity. *Proc Natl Acad Sci* 106(26):10570–10575
- Hidalgo CA, Klinger B, Barabási AL, Hausmann R (2007) The product space conditions the development of nations. *Science* 317(5837):482–487
- Horbach J, Rammer C, Rennings K (2012) Determinants of eco-innovations by type of environmental impact: the role of regulatory push/pull, technology push and market pull. *Ecol Econ* 78:112–122
- Hufbauer GC, Kim J (2010) Reaching a global agreement on climate change: what are the obstacles? *Asian Econ Policy Rev* 5(1):39–58
- Jaffe AB, Palmer K (1997) Environmental regulation and innovation: a panel data study. *Rev Econ Stat* 79(4):610–619
- Johnson B, Lorenz E, Lundvall BÅ (2002) Why all this fuss about codified and tacit knowledge? *Ind Corp Chang* 11(2):245–262
- Laursen K (1998) Revealed comparative advantage and the alternatives as measures of international specialisation (No. 98-30). DRUID, Copenhagen Business School, Department of Industrial Economics and Strategy—Aalborg University, Department of Business Studies
- Liu J, Goldstein D (2013) Understanding China's renewable energy technology exports. *Energy Policy* 52:417–428
- Marin G, Vona F (2019) Climate policies and skill-biased employment dynamics: evidence from EU countries. *J Environ Econ Manag* 98:102253
- McMillan M, Rodrik D, Verdusco-Gallo I (2014) globalization, structural change and productivity growth, with an update on Africa. *World Dev* 63:11–32

- Mealy P, Teytelboym A (2020) Economic complexity and the green economy. *Res Policy* 51:103948
- Melitz MJ (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71(6):1695–1725
- Merino-Saum A, Clement J, Wyss R, Baldi MG (2020) Unpacking the green economy concept: a quantitative analysis of 140 definitions. *J Clean Prod* 242:118339
- Midelfart-Knarvik KH, Overman HG, Redding SJ, Venables AJ (2000) The location of European industry. In: *Economic Papers* 142, European Communities
- Moretti E (2010) Local multipliers. *Am Econ Rev* 100(2):373–377
- Mulatu A, Gerlagh R, Rigby D, Wossink A (2010) Environmental regulation and industry location in Europe. *Environ Resour Econ* 45:459–479
- Nesta L, Vona F, Nicolli F (2014) Environmental policies, competition and innovation in renewable energy. *J Environ Econ Manag* 67(3):396–411
- Nesta L, Verdolini E, Vona F (2018) Threshold policy effects and directed technical change in energy innovation. GREDEG working paper 01-2018
- Nicoletti G, von Rueden C, Andrews D (2020) Digital technology diffusion: A matter of capabilities, incentives or both?. *Eur Econ Rev*, 128:103513
- Nordhaus W (2019) Climate change: the ultimate challenge for economics. *Am Econ Rev* 109(6):1991–2014
- Nordhaus WD, Boyer J (2000) *Warming the world: economic models of global warming*. MIT press
- Perruchas F, Consoli D, Barbieri N (2020) Specialisation, diversification and the ladder of green technology development. *Res Policy* 49(3):103922
- Petralia S, Balland PA, Morrison A (2017) Climbing the ladder of technological development. *Res Policy* 46(5):956–969
- Popp D (2002) Induced innovation and energy prices. *Am Econ Rev* 92(1):160–180
- Popp D, Vona F, Marin G, Chen Z (2021) The employment impact of green fiscal push: evidence from the American Recovery and Reinvestment Act. *Brookings papers on economic activities*
- Rajan R, Zingales L (1998) Financial development and growth. *Am Econ Rev* 88(3):559–586
- Requate T (2005) Dynamic incentives by environmental policy instruments: a survey. *Ecol Econ* 54(2–3):175–195
- Rodrigues JF, Moran D, Wood R, Behrens P (2018) Uncertainty of consumption-based carbon accounts. *Environ Sci Technol* 52(13):7577–7586
- Romalis J (2004) Factor proportions and the structure of commodity trade. *Am Econ Rev* 94(1):67–97
- Romer PM (1990) Endogenous technological change. *J Political Econ* 98(5, Part 2):71–102
- Sato M (2014) Product level embodied carbon flows in bilateral trade. *Ecol Econ* 105:106–117
- Sauvage J (2014) The stringency of environmental regulations and trade in environmental goods, OECD Trade and Environment Working Papers, OECD Publishing, Paris
- Sawhney A, Kahn ME (2012) Understanding cross-national trends in high-tech renewable power equipment exports to the United States. *Energy Policy* 46:308–318
- Sbardella A, Perruchas F, Napolitano L, Barbieri N, Consoli D (2018) Green technology fitness. *Entropy* 20(10): 776.
- Shapira P, Gök A, Klochikhin E, Sensier M (2014) Probing “green” industry enterprises in the UK: a new identification approach. *Technol Forecast Soc Chang* 85:93–104
- Soete LLG, Verspagen B (1994) Competing for growth: the dynamics of technology gaps. In: Pasinetti L, Solow R (eds) *Economic growth and the structure of long-term development*. Macmillan, London
- Steenblik R (2005) Environmental goods: a comparison of the APEC and OECD lists (No. 2005/4). OECD Publishing, Paris
- Szirmai A, Verspagen B (2015) Manufacturing and economic growth in developing countries, 1950–2005. *Struct Chang Econ Dyn* 34:46–59
- Tamini LD, Sorgho Z (2018) Trade in environmental goods: evidences from an analysis using elasticities of trade costs. *Environ Resource Econ* 70(1):53–75
- Van Beveren I, Bernard AB, Vandenbussche H (2012) Concording EU trade and production data over time (No. 18604). National Bureau of Economic Research
- Vona F, Marin G, Consoli D, Popp D (2018) Environmental regulation and green skills: an empirical exploration. *J Assoc Environ Resour Econ* 5(4):713–753
- Vona F, Marin G, Consoli D (2019) Measures, drivers and effects of green employment: evidence from US local labor markets, 2006–2014. *J Econ Geography* 19(5):1021–1048
- Weitzman ML (1998) Recombinant growth. *Q J Econ* 113(2):331–360
- Wiebe KS, Yamano N (2016) Estimating CO<sub>2</sub> emissions embodied in final demand and trade using the OECD ICIO 2015. OECD Publishing, Paris
- World Bank (2007) *International trade and climate change*. The World Bank, Washington

- World Bank (2012) Inclusive green growth the pathway to sustainable development. The World Bank, Washington
- WTO (2001) Doha Ministerial Declaration, Ministerial Declaration Adopted on 14 November 2001
- Yu R, Cai J, Leung P (2009) The normalized revealed comparative advantage index. *Ann Reg Sci* 43(1):267–282
- Zugravu-Soilita N (2018) The impact of trade in environmental goods on pollution: what are we learning from the transition economies' experience? *Environ Econ Policy Stud* 20(4):785–827
- Zugravu-Soilita N (2019) Trade in environmental goods and air pollution: a mediation analysis to estimate total, direct and indirect effects. *Environ Resource Econ* 74(3):1125–1162

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.