transeuroworks

Working Paper 2/2025

Deliverable 2.4.

The green employment index: a composite indicator to assess the greenness of employment in the EU

Alexandre Ounnas

transeuroworks

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. N. 101061198

INTRODUCTION

The green transition, once largely a topic of discussion, has seen considerable policy action in Europe in recent years, perhaps most notably culminating in the adoption of the European Green Deal in 2020. Yet, the general high level of public support (at least in the recent past, European Commission, 2021), has come with pushbacks (e.g. the Yellow Vests movements in France), demonstrating the need for policymakers to approach the green transition from a socioeconomic perspective, in addition to a technological and environmental one.

Indeed, as European economies – and labour markets – becomes greener in pursuit of a sustainable economy, the socioeconomic implications are likely to be substantial, with some even equating it a modern-day industrial revolution (Bowen et al. 2016). The concept of 'green job' then becomes central for policymakers to understand and anticipate the likely socioeconomic impacts associated with the green transition.

No agreed-upon definition of a green job exists at international level, though a relative consensus seems to emerge around the definition developed by the Green Job Initiative (Stanef-Puică et al. 2022). Green jobs are defined as those '*that contribute substantially to preserving or restoring environmental quality*' (UNEP/ILO/IOE/ITUC, 2008). A list of associated green activities is provided as well, but this definition is broad enough that it offers a certain leeway in the identification of jobs fitting under its umbrella.

It is therefore not surprising to see emerge different methodologies drawing on different approaches, as regard for example, sectors, processes, output, to measure green jobs (Bowen et al. 2018; Vona et al. 2018; Bontadini & Vona, 2023). This variety of approaches results in a wide range of estimates for the number of green jobs expressed as a share of total employment – from 0.7% to 19.4% of total employment in the US – ultimately undermining important policy work (Bowen & Kuralbayeva, 2015; Apostel & Barslund, 2024).

This difficulty can be seen as arising from the difficult task of summarising a complex multidimensional concept in a binary outcome. Recent contributions from academics and (inter)national institutions appear to acknowledge this complexity by considering different approaches jointly for measuring and analysing green jobs (Bontadini & Vona, 2023; OECD, 2024). Along this line, this study presents an attempt at developing a composite indicator integrating the main approaches used in the literature to measure the green jobs.

Composite indicators have been produced by many institutions around the world, and are used for evaluating and tracking progress in different policy areas (Nardo et al. 2005). These indicators can be useful for policymakers as they synthetise in one single index value, the information conveyed by a multitude of indicators. In recent times, the increased amount of available data appears to have raised the popularity of these indicators (Greco et al. 2019), although concerns and critics are often heard around their limitations, whether methodological or with regard to the interpretation of the results.

Provided that care is taken at each step of the construction process (Nardo et al. 2005), a composite index can nonetheless prove to be a useful tool for the study and measure of complex multi-dimensional phenomena. In the literature, no attempts at constructing a green job composite indicator have been identified, whereas the variety of existing approaches suggest potential gains from integrating these methodologies under a single framework.

This study thus proposes an attempt at constructing a green job index (GJI) for EU Member States (MS). This index aims to evaluate the greenness of a job by aggregating information from several indicators.

The detailed methodology proposed by Nardo et al. (2005) for constructing sound composite indicators is used to guide the decisions made at different steps of the process. The framework further builds on existing composite indicators, the Gender Equality Index and the Industrial Relation Index in particular (Caprile et al. 2017; European Institute for Gender Equality, 2017), and organise its structure around dimensions, subdimensions and indicators. The GJI is available at rather fine level[1] and is obtained from 25 indicators, arranged in 8 subdimensions and 5 dimensions.

Some indicators are standards and include for instance, task, air emissions, value added in the Environmental Good and Service Sector (EGSS). Other indicators related to energy use or waste generation are seldom used, though they bear a link with climate goals. All indicators are selected from publicly available sources (primarily Eurostat) to ensure quality, comparability and timeliness of the indicators.

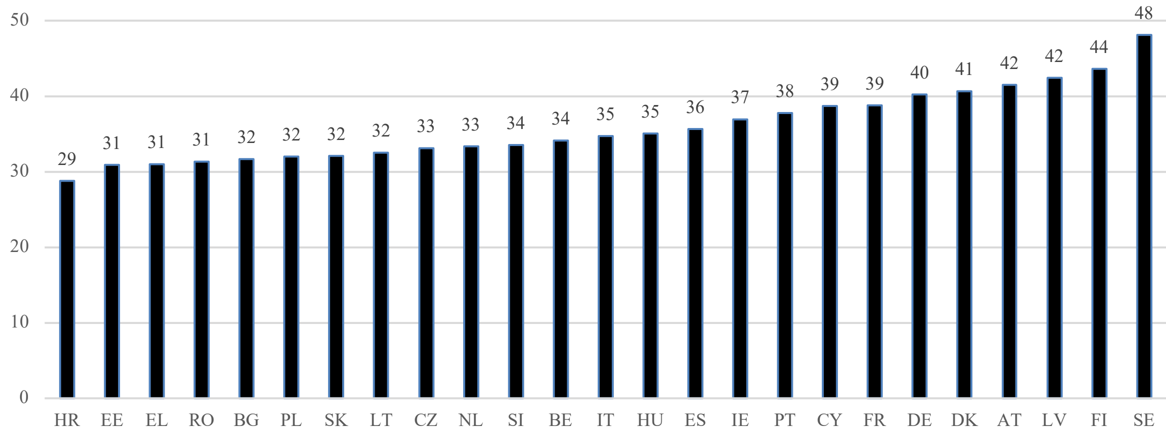
[1] An index value is available for each job at NACE 1-digit sector and ISCO08 3-digit occupation. This represent the finest level of information on sectors and occupations provided in anonymised files of the European Labour Force Survey.

The GJI can be used to identify green jobs through different approaches – e.g. by setting a threshold value above which a job is considered green –. However, transforming the GJI in a binary outcome green/non green might not represent the best possible use for the index. Rather, the continuous nature of the index is preserved and used to evaluate the greenness of employment in MS and its evolution through time. This achieved by matching the GJI to the European Labour Force Survey (EU LFS). The Green Employment Index (GEI) can then be computed using employment shares obtained from the EU LFS as weights.

The GEI is higher in MS with better environmental performances as measured through the selected indicators (i.e. low emissions, low generation of waste). It measures the greenness of aggregate employment rather the share of green jobs, by exploiting to continuous scores it provides. The GEI can be used to track progress in the greening of employment across MS and time.

First versions of the GJI and GEI have been produced for this paper. However, results should be considered with care given that the construction of a composite requires many sensitivity and robustness checks, that could not be performed for these versions of the indices.

Figure 1: Green Employment Index – 2022



Source: Own elaboration.

Bearing in mind these limits, employment in 2022 was the greenest in SE with the GEI value of 48, followed by FI, LV and AT. Employment was the least green in HR, EE and EL. Analysis of the GEI evolution between 2011 and 2022 reveals that employment became greener in all but one MS. The GEI increased substantially in DK, but also in many MS with initial GEI values below average. This suggest a certain convergence in the greenness of employment over the period, although differences remain important in 2022.

The rest of the paper is organised as follows. Section 1 reviews some of the existing literature of the literature that will serve to inform the construction of the GJI and GEI. Sections 2 focuses on the construction of the GJI. This section includes a discussion of the conceptual framework, the selection of indicators, the quantitative methodology and the results for this index. Section 3 concludes with the presentation of the GEI and some early results.

1. Literature review

1.1. Literature on Green Jobs and Employment

This section reviews the concept of green jobs used in the academic and grey literature. This literature review is not meant to be exhaustive but aims at identifying the main methodological approaches used to characterise and measure green jobs. More detailed literature reviews on the concept can be found in Stanef-Puică et al. (2022) or Apostel & Barslund (2024).

As frequently noted in the literature (Bowen & Kuralbayeva, 2015; OECD, 2024; Urban et al. 2023), there is no internationally agreed-upon definition of what constitutes a green job. Nevertheless, it is possible to identify a relative consensus around the idea that green jobs contribute to preserve and restore the environment, and mitigate climate change and the adverse effects of human activities on environmental quality (Stanef-Puică et al. 2022). The definition initially developed by the joint UNEP, ILO, IOE, ITUC green jobs initiative (UNEP/ILO/IOE/ITUC, 2008) can serve as reference and defines green jobs as '*as work … that contribute substantially to preserving or restoring environmental quality. … this includes jobs that help to protect ecosystems and biodiversity; reduce energy, materials, and water consumption through high-efficiency strategies; de-carbonize the economy; and minimize or altogether avoid generation of all forms of waste and pollution.*'

This broad definition can encompass jobs in traditional sectors and occupations but also jobs in new and emerging activities[2]. As a result, it is not surprising to see a variety of approaches emerge to

[2] For instance, the Green New and Emerging Occupations identified by Dierdorff et al. (2009).

'operationalise' the definition and measure the extent of green jobs[3]. Ultimately, this heterogeneity reflects the multidimensional nature of the green job concept and its complexity.

The main methodological approaches can be classified broadly as those taking a sectoral perspective focused on production, and those seeking to identify green jobs based on occupations and the content of the job (Bluedorn et al. 2023)[4]. Furthermore, these methodologies are often split between output and processes (Vona et al. 2018; Maldonado et al. 2024). This categorisation is used below to organise the presentation of the different methodologies found in the literature.

It is also worth mentioning that alternative approaches (e.g. the systemic approach, Bohnenberger, 2022) exist in the literature, though these are generally complicated to implement from a data point of view and are not discussed in this review. See Urban et al. (2023) for a discussion of these alternatives.

1.1.1 Occupations

Occupational approaches classify green jobs based on the content of the job performed rather than the industry in which the job takes place. Tasks and skills constitute the two main approaches to measure green jobs in this context. The two terms are often used interchangeably in the literature, although they refer to different concepts (Acemoglu & Autor, 2011).

Green tasks and green skills

The task-based approach is well-known in the academic literature, as it has been used extensively to identify and study the job polarisation phenomenon (Goos & Manning, 2007; Autor & Dorn, 2013). In the context of green jobs, a first reference to tasks emerged from the work of Dierdorff et al. (2009), who classified green jobs in three categories, including green enhanced skills and new and emerging jobs, two categories for which the development of the green economy would alter the content of the job.

---

[3] A link can be made with the concept of sustainability where there is a broad agreement on its definition (from the Brundtland report) but no '*consensus regarding the operational content of environmentally sustainable economic development*' (Mori & Christodoulou, 2012).

[4] Bluedorn et al. (2023) refer to what workers do (the occupation) and where they work (the sector).

Relying on the work from the Green economy initiative of the Occupational Information Network (O*NET), Vona et al. (2018) exploited information on tasks labelled as green to construct an indicator consisting in the share of green tasks per occupation. The task approach has been used numerous time to quantify the extent of green jobs, proving to be a useful and popular approach for achieving this purpose (Valero et al. 2021; Vona, 2021; Maldonado et al. 2024).

To identify green jobs from the task indicator, an additional step is required to determine whether green jobs corresponds to those with at least one green task ;(Bowen & Hancké, 2019), or those with a more substantial green task content (Vona et al. 2018).

Comparatively, there are less contributions relying on green skills to identify green jobs, and the concept has appeared to be more popular amid international organisations such as the ILO and CEDEFOP, which have developed green skills frameworks to map workforce capabilities in sustainability-related fields.

More recently, contributions have started to explore the possibility offered by the European Skills, Competences, Qualifications and Occupations (ESCO) classification. ESCO provides information on more than 13 000 skills and knowledges, 570 of them being labelled as green[5]. Recently, Maldonado et al. (2024) used ESCO to construct an indicator equivalent to the share of green skills per occupation, and ANPAL, (2023) made use of the database to construct different green skills indices by occupation.

Occupational approaches to measure green jobs are generally part of the methodologies used by several national institutions, including the 'Observatoire National des Emplois et Métiers de l'économie Verte' in France or the Office for National Statistics in the U.K.

1.1.2 Sectors

The sector approach identifies green jobs from the output and processes associated with production.

Output

Output-Based approaches classify green jobs based on the production of environmental goods and services.

---

[5] The ESCO database is discussed further in Annex B1.1.

The Environmental Goods and Services Sector (EGSS) framework, developed by Eurostat (2016), identifies jobs in industries that produce goods or services for environmental protection or sustainable resource management. This approach is widely followed in EU statistical reporting and is often used in the literature to identify green jobs as those that can directly be related to the production of green output. For instance, this approach has been employed by Valero et al. (2021), Maldonado et al. (2024); and OECD, (2024).

Other output-based approach include the (now discontinued) green goods and services survey ran by the U.S. Bureau of Labor Statistics, which classified green jobs based on the proportion of industry revenue derived from green products (Sommers, 2013). In addition to output from EGSS, the Office for National Statistics in the U.K makes use of the Low Carbon and Renewable Energy Economy survey to complement its estimate based on jobs in green industries.

Process

Process-based approaches classify green jobs by relying on the environmental performance of sectors or industries. This approach tends to broaden the pool of potential green jobs since the greenness of jobs becomes solely determined by its location, and is no longer associated with the (rather niche) production of green output (OECD, 2024).

This approach is often used in conjunction with the task or output approach to analyse their overlap (Bontadini and Vona, 2023). Indicators are usually constructed from Greenhouse gas (GHG) emission intensities (or other air emissions) obtained from emission inventories (OECD, 2024) or from the carbon footprints of industries retrieved through Input-Output analysis (Bontadini and Vona, 2023).

Similarly to the task approach, methodologies based on process indicators require setting a threshold to identify low emission sectors. This is usually achieved through the identification of a group of best performers and any job part of low emission sector is then considered green.

Though not mentioned in this brief review, these methodological approaches can suffer from limitations and shortcomings, often related to data availability and aggregation (for the task approach, see for instance Vona, 2021 and Bachelot, 2024). These are further discussed in Section 2 and Annex B1.

In the following sections, I develop a framework to construct a green job composite indicator, which allows for integrating these various approaches and obtain a single index value informing on the greenness of a job. The next section shortly reviews the literature on composite indicators.

1.2. Composite indicators

The construction of composite indicators is central to policymaking and many international organisations produce and rely on these indicators (e.g. the Human Development Index) to evaluate and track progress in different policy areas (Nardo et al. 2005). There is no official definition of what constitute a composite indicator (Greco et al. 2019), although the different definitions gather around the idea of studying and measuring complex multi-dimensional phenomenon by aggregating series, often expressed in different units, into a single index value. Composite indicators are used in different fields including economic and social analysis such as business cycles analysis (Mariano & Murasawa, 2003; Stock & Watson, 2008), innovations (Dutta et al. 2022), industrial relations (Kim et al. 2015; Caprile et al. 2017), well-being (World Health Organization., 2024), and job quality (Leschke & Watt, 2014; Cascales Mira, 2021). Composite indicators have also been used to study concepts related to energy (Nussbaumer et al. 2013), the environment (Juwana et al. 2012; Hsu & Zomer, 2016) and many other areas relevant to policymakers. Furthermore, some composite indicators can be seen as transversal as they study concepts across different fields. This is typically the case of sustainability indicators which aggregate social, economic and environmental indicators (Mori & Christodoulou, 2012; Madurai Elavarasan et al. 2022).

In addition to synthetising information from multiple indicators, a composite index allows for comparisons between units (e.g. countries) and through time (Nardo et al. 2005). These indicators have become increasingly popular over the last decades, spurred by the increased amount of available data (Greco et al. 2019). However, these indicators have also been extensively criticised as they suffer from weaknesses. Whilst very useful to convey information in a simplified manner, the aggregation of many series can lead to misleading conclusions. More importantly, there are several steps to follow for constructing such indices and many involves non-trivial decisions that can have important consequences for the final index value. Core issues are usually associated with weighting and aggregation of the individual indicators (Greco et al. 2019)[6]. It is therefore paramount to be as transparent as possible throughout the construction of the composite index and perform robustness checks and sensitivity analysis to understand the impact of the different modelling decisions.

These critics have been (partly) answered through the development of a substantial literature on methodological considerations and best practices for the construction of sound indicators. Nardo et al. (2005) developed a methodology which can be considered a reference and aims at ensuring the construction of transparent and meaningful composite indicators. The proposed methodology consists in 10 steps (displayed in Annex A) and is used as a guide for the construction of the green job index (GJI) in the following sections.

This methodology has been applied to develop several indices including the gender equality index of the European Institute for Gender Equality (European Institute for Gender Equality., 2017) and Eurofound’s industrial relations index (Caprile et al. 2017). These two composite indicators are particularly interesting for the construction of their conceptual frameworks, which includes a hierarchical structure that can help understand the impacts of each individual indicator on the composite index. These indices are structured such that the concept of interest (e.g. gender equality) is split into dimensions (or domains), themselves composed of subdimensions associated with indicators. The gender equality index is obtained from 31 indicators aggregated into 14 subdimensions and 6 dimensions (European Institute for Gender Equality, 2017; Papadimitriou et al, 2020) and the industrial relations index includes 4 domains, 16 subdomains and 45 indicators (Caprile et al 2017). These two indices are important sources of information used as references for the construction of the GJI.

2. A COMPOSITE INDICATOR TO MEASURE THE GREENNESS OF JOBS

The literature review in the previous section has underlined the variety of methodologies used to measure green jobs.

Currently, the dominant approach is probably to measure green jobs using the task content of occupations (Dierdorff et al. 2009; Vona et al. 2018). However, this approach is likely leading to a

[6] A particular issue relates to the use of aggregation methods that result in compensation between domains and sub-domains of the composite indicator (Nardo et al. 2005; Munda, 2012).

narrow measure of jobs that could potentially fall under the definition of green jobs. For instance, it is well-admitted that many jobs important for the green transition (i.e. climate adaptation and mitigation) are already existing and will not see their tasks content substantially affected (Dierdorff et al. 2009; Maldonado et al. 2024)[7]. Some of these jobs and tasks can also be linked to the circular economy where maintenance, repair for second use, and recycling are important skills/tasks partially captured by the occupation approach.

Similarly, the mapping between green jobs and green output (as measured through EGSS) is imperfect since output is measured at sectoral level. As a result, it is likely that occupations without any green tasks/skills contribute to produce green output and green occupations can also take place in sector that produce non-green output.

Hence, both the content of the job as well as where the job takes place (sector) appear as relevant dimensions to characterise the greenness of a job. These two dimensions only partially overlap and as such, they could be considered jointly to ensure that the different characteristics of green jobs they identify are considered in their entirety. Furthermore, it is important to note that the sector dimension is considered beyond green output to include aspect related to (production) processes (e.g. air emissions). As explained in Section 1.1.2, the process dimension increases the pool of potential green jobs as those taking place in more environmentally friendly sectors, without necessarily producing green output, could also be considered green.

Overall, these observations point to a potential benefit from integrating these different approaches to try capturing the concept of green jobs in a more complete way. The composite indicator approach is relevant in that it acknowledges the information in each different approaches, offering a potential way around some of the limitations appearing when these approaches are considered individually.

It should be noted that, to some extent, these shortcomings are already recognised in the literature since most of the cited contributions rarely rely on one single approach to measure green jobs. The main advantage of the composite indicator is then to provide an organised framework to jointly consider all these different methodologies and avoid certain issues (e.g. double counting) arising when methodologies are not integrated together.

[7] These are captured, to some extent, by the group of Green Increased Demand Occupations, defined by Dierdorff et al., (2009). Maldonado et al. (2024) further mention the example of public bus drivers.

The construction of the composite indicator aims to assign a score measuring the greenness of a job. The index should be able to rank and compare MS through time which imposes restrictions (i.e. in terms of representativeness and harmonisation) on the data available, and hence on the level at which a job can be measured. A job can be defined as '*a set of tasks and duties performed, or meant to be performed by one person for a single economic unit*' (Stoevska & Hunter, 2012) and occupations are themselves bundles of jobs with similar characteristics. The EU-LFS provides information on jobs at 3-digit ISCO-08 occupation level, a rather aggregated level. To increase the granularity of the data, 3-digit occupations can be interacted with information on NACE 1-digit sectors, with the underlying idea that information on sectors could potentially help pinpoint different jobs falling under the same 3-digit occupation. Furthermore, this definition provides a natural way to integrate sectoral indicators in the measure of green jobs and will avoid issues related to double counting of jobs that could be considered green based on different indicators.

Moreover, the score obtained from this green job index (GJI) would not provide a direct way to determine which jobs are green. To avoid imposing an ad-hoc threshold is used assess the overall greenness of aggregate employment, through the computation of the green employment index (GEI). The GEI can be seen as the final output from this study and will assign higher values to MS with higher shares of employment in greener (rather than green) jobs, as measured through the GJI.

Hence, the GJI constitutes an intermediate step for the computation of the final GEI. The work is therefore divided in two steps. The first step will produce the GJI for each job to measure its greenness through the aggregation of various indicators related to e.g. energy, waste generation, air emissions, tasks. In the second step, the scores for each job will be merged with EU-LFS data and the final GEI will be computed by aggregating the GJI values using employment as weights. Ultimately, the final GEI will inform on the overall greenness of aggregate employment in each MS and its evolution through time without requiring the determination of which jobs are green.

The composite indicator to score the greenness of jobs (first step) relies on the methodology of Nardo et al. (2005) displayed in Annex A. The construction of this composite indicator is presented in this section, starting with the definition of the conceptual framework. The construction of the GEI using EU-LFS data is explained in Section 3.

2.1. An integrated conceptual framework

The conceptual framework supporting the construction of the green job index (GJI) relies on the UNEP, ILO, IOE, ITUC definition recognising that the ultimate goal of green jobs are to contribute to the protection, preservation or restoration of the environment and to limit the environmental impact of human activities (see Section 1.1). The willingness to integrate different approaches found in the literature implicitly acknowledges that these goals can be achieved in different ways and various indicators, most of them already used in the literature, can be linked to these goals.

Following the literature review discussed in Section 1, the conceptual framework, displayed in Figure 2, is organised around two pillars, namely the sector (or production) pillar and the occupation (or ob content) pillar. Each of these pillars is decomposed following the same hierarchical structure used for the constructions of the gender equality index and industrial relations index (European Institute for Gender Equality. 2017; Caprile et al. 2017) with domains (or dimensions), subdomains and indicators. These two pillars are aggregated together at the end of the construction process to obtain a single index value measuring the greenness of each single job, understood as the combination of a 3-digit ISCO-08 code and a 1-digit NACE code .

An important observation from the literature is that indicators are usually classified into process and output indicators. This categorisation in (input,) process and output constitute a common approach to analyse various (social) concepts (Kim et al. 2015) and is acknowledged by Nardo et al. (2005) as a possible way to identify dimensions and sub-dimensions for the construction of a composite indicator. I rely on this system approach and the production and job content pillars are both divided into an input, a process and an output dimension. This categorisation has the advantage of clarifying the associations between dimensions since inputs are used as part of processes to create output. This implies that the dimensions making the composite index should be correlated at least to some extent[8].

Subdimensions can then be identified for each of the input/process/output dimensions. For output, a single subdimension is used, simply defined as green production and measured through e.g. EGSS

[8] For instance, the move away from fossil fuel as an energy input should lead to lower GHG emissions from the production process.

and outputs from eco-innovations (EC: Directorate-General for Research and Innovation, 2024). The occupation pillar does not feature an output dimension given that output is measured at sectoral level[9].

The process dimension is usually associated with subdimensions related to green tasks (occupation pillar) and air emissions (e.g. GHG emission) for the sector pillar. Waste generation is an additional relevant subdimension for the process dimension in the sector pillar given its environmental impact and its relation to the circular economy.

In general, the input dimension has been disregarded in the green job literature. However, this dimension is relevant for the definition of green jobs as the associated indicators also inform on whether jobs take place in sectors contributing to environment goals. For instance, data on land use is relevant with regards to environment preservation and energy use as an input to the production system is directly linked to climate adaptation and mitigation goals. To identify subdimensions for inputs, I think of inputs entering a production function. These are usually labour and capital and I further consider energy and land.

The last three subdimensions (i.e. capital, energy and land) are naturally assigned to the sector pillar, while the labour subdimension is included as an input in the occupation pillar. The indicators for the labour subdimension will rely on (green) skills, an information provided at the occupational level.

It is worth noting that skills and tasks enter the composite indicator through separate (sub)dimensions. This distinction could be superficial given that there exists a (imperfect) mapping between skills and tasks[10]. However, it is conceptually correct and motivated by the theoretical framework given that skills, a worker's attributes, are applied to tasks (process) to generate output (Acemoglu & Autor, 2011). Moreover, it could be important to keep skills and tasks initially separated, as from a policy

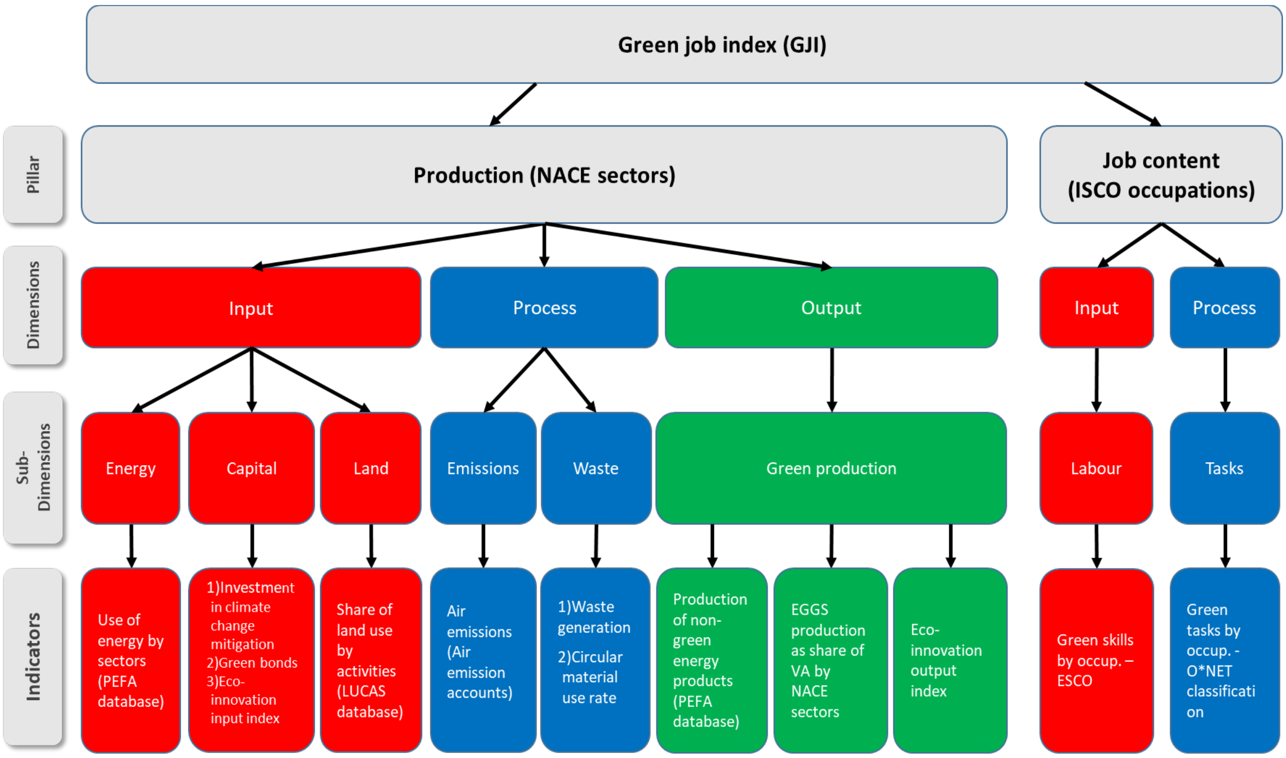
[9] In Urban et al., (2023), the conceptual framework included a job quality subdimension under the output dimension. This dimension has been dropped in the current version, mainly for simplifying the framework.

[10] When considering green jobs, this mapping could be even less obvious as for instance, Maldonado et al. (2024) report that only four occupations are considered green according to both the task and the skill approach. In this study, the correlation between the three skill indicators and the task indicator ranges between 0.6 and 0.7. This evidence highlights this imperfect mapping, although differences in the methodologies to identify green skills and green tasks could also explain this imperfection.

standpoint, skill acquisition is addressed through education and training, while tasks are related to the organisation of work within firms[11].

The conceptual framework displayed in Figure 2 highlights the hierarchical structure of the GJI and its integrated nature as for instance, the EGSS and task approach are included in the computation of the composite indicator. Moreover, an important advantage of this structure is that index values will be created (based on the indicators) for each subdimension, which will then be aggregated into dimensions up to the final GJI. This will allow for tracking whether certain dimensions drive the results obtained for individual scores. Furthermore, the (sub)dimensions can be analysed individually to focus on specific characteristics of green jobs.

Figure 2: Conceptual framework



Source: Own elaboration

[11] This point can be linked to the concept of greening (Castillo, 2023). In this context, rethinking tasks and more broadly the organisation of work, could be important lever to foster greening at the workplace. Furthermore, this could be achieved with none or small modifications in the skill set required to perform jobs. For instance, the promotion of certain practices in agriculture (e.g. compost, bio) would modify tasks without deeply modifying the skills required for the job.

2.2. Data collection

The selection of indicators constitute the second step of the methodology proposed by Nardo et al. (2005). This selection is very important and should be informed by the conceptual framework and driven by different considerations. These include the availability and quality of the data, as well as the scope and aim of the composite indicator. For the GJI, availability of data for all (or most) MS at a relatively high time frequency, ideally yearly, constitute central considerations to ensure that the index can be used for comparisons between MS through time.

For these reasons, the data selection is mainly restricted to series produced by Eurostat under the 'environment' categorisation, which should further guarantee timeliness and quality of the data. Eurostat series are used only for the sector pillar since environmental data is (often) provided at 1-digit NACE level, though service sectors are usually aggregated into one broad sector. These series are complemented with two series retrieved from the European Environment Agency. This data is not available at sectoral level (only at MS level), but the quantitative methodology and the aggregation of indicators (discussed in Section 2.3) will account for this discrepancy.

Data for the occupation pillar is retrieved from O*Net for tasks and ESCO for skills. O*Net is the usual source of information regarding green tasks (Consoli et al. 2015; Vona et al. 2018) and ESCO has developed a categorisation that identifies more than 500 green skills (ESCO, 2022), which can be used to generate an indicator of green skills by occupations (see e.g. ANPAL 2023).

In total, 25 indicators are associated with the 8 identified subdimensions (e.g. Energy, Capital). These indicators are always expressed such that an increase in their values is consistent with a greener outcome[12]. It is worth emphasizing that the use of sectoral data has the major advantage of introducing some variability across MS and time (something not offered by occupation data beyond variations in the occupational composition of employment). As a result, each indicator is available

[12] For instance, air emissions are usually expressed in terms of value added as air emissions intensities. Greening would then be associated with a decrease in the indicator. In this study, the indicator related to air emissions will be constructed as the inverse of intensities by dividing VA by air emission intensities. Though not necessary, this decision simplifies the construction of the composite indicator. For instance, this transformation is implemented for the construction of the gender equality index (European Institute for Gender Equality, 2017).

as multiple series (i.e. by MS and sectors) and the composite indicator ultimately includes close to 5000 series in total.

For the energy subdimension, Eurostat's data on energy flows from the Physical Energy Flow Account is used to create five indicators namely, energy productivity using sectoral value added (VA), the share of (non-)green energy products used as inputs and the share of (non-)renewable natural resources used to produce energy products. The data does not provide information on the source of electricity and this energy product is therefore not used to create indicators.

The capital subdimension includes three indicators. The first one makes use of data form the Environmental Protection Expenditure Accounts on investment in climate mitigations expressed in terms of sectoral VA. The other two indicators are retrieved from the European Environment Agency (EEA) and are the share of green bonds and the eco-innovation input index (EC: Directorate-General for Research and Innovation, 2024).

The land subdimension relies on land use data by 1-digit NACE sectors available from the Land Use and Coverage Area Frame Survey (LUCAS). Data on land use is collected every three years and is currently available between 2009 and 2018 only (four data points).

Regarding the Labour-input subdimension, the European Skills, Competences, Qualifications, and Occupations (ESCO) database is used to construct three indicators. ESCO provides information on more than 13,000 skills and knowledges (or concepts) assigned across 3,000 occupations. The ESCO occupation classification expands on the 4-digit ISCO classification and it is therefore required to aggregate the data back to the 3-digit ISCO level. The different adjustments to the ESCO data are briefly discussed in Annex B1.2

ESCO provides its own classification of green concepts defined as 'the knowledge, abilities, values, and attitudes required to support environmental sustainability across occupations and industries' (ESCO, 2022). In total, there are 570 concepts labelled as green. Skill and knowledge are further classified into essential and non-essential, where non-essential should be understood as skill relevant only for a subset of jobs within the occupation (ESCO, 2018).

The first indicator consists in the share of green concepts among the total number of concepts per occupation. Two additional indicators are constructed from the share of essential skills and essential knowledges expressed in terms of the total number of skills/knowledges. The final indicators used to compute the skill subdimension are further transformed as $log(1 + \text{the indicator})$ since the share of green skills/knowledges is large for a couple of 3-digit occupations (e.g. 621 – Forestry and Related Workers and 961 – refuse workers) and this transformation allows for 'smoothing' larger values (see Figure 47).

Table 1 displays a list of these indicators, which are further discussed in Annex B1.

Table 1: Input indicators

Subdimension	Source	Series name	Series code	brief description	indicators	remarks
Energy	Eurostat	Energy supply and use by NACE Rev. 2 activity	env_ac_pefasu	Data originates from Physical Energy Flow Accounts (PEFA). PEFA collects data on energy flows from (1) the environment to the economy; (2) within the economy (e.g. intermediate input, final consumption); (3) back to the environment from the economy.	1) energy productivity 2) share of green energy products in total energy used as input 3) share of non-green energy products 4) share of renewable natural resources 5) share of non-renewable natural resources	missing for EL – L - missing for SE – B, C and D MS level MS level
Capital	EEA	Green bond issuance by corporates and governments	EU SDG 13_70a	Bonds used to finance activities that address climate change and environmental issues - green bonds - provide a means to increase green investment	1) percentage of green bonds as a share of the total (corporate and governments) number of bonds issued.	MS level
Capital	EEA	Eco-innovation index	SUSO001	Eco-innovation refers to any innovation that reduces impacts on the environment, increases resilience to environmental pressures or uses natural resources more efficiently. The eco-innovation index is based on the eco-innovation scoreboard, which has indicators in five thematic areas	1) Index on eco-innovation inputs, which includes financial and human capital investment in eco-innovative activities.	MS level
Capital	Eurostat	Investments in climate change mitigation by NACE Rev. 2 activity	env_ac_ccminv	Investment in climate change mitigation defined as investment in sectors economic goods and services that substantially reduce GHG emissions.	1) Investment as a share of sectoral VA	Available for all MS only for sectors C, D, F, H, J and M
Land	Eurostat	Land use	lan_use_ovw	Data is collected from LUCA survey ran every 3 years. The dataset provides information on land cover (grassland) and land use (e.g. private gardens, public parks, agriculture)	1) Land use by sector a share of total available land	Missing for: EE – E, F, G, I, K, L, M, N, O, P, Q, S; HR – all sectors; SI – G, I, K, L, M, N, O, P, Q
Labour	DG EMPL	ESCO	-	The ESCO database provides information on the skills and knowledge by occupations. A green skills/knowledge classification is available.	1) Share of green skills/knowledges by occupation 2) Share of essential green skills by occupation 3) Share of essential green knowledges by occupation	data missing for occupation 224, 631 and 632

Source: Own elaboration based on available information and the relevant literature.
Note: See Annex B1.1 and B1.2 for additional details on the datasets and indicators. The NACE and ISCO classification can be accessed from https://en.wikipedia.org/wiki/NACE_rev2 and https://ilostat.ilo.org/classification-occupation/.

Regarding the Waste subdimension, some indicators can be constructed from Waste Generation Statistics, which provide comprehensive data on the types and quantities of waste produced across the European Union. Data is available at sectoral level and offers a detailed breakdown by waste categories for more than 30 material types (e.g. metals, plastics). Data on waste generation is used to construct four indicators. These include series on VA per one tonne of waste for each available sectors, and the share of recyclable, chemical and hazardous wastes among total sectoral wastes. Waste generation statistics are collected on bi-annual basis and the service sector aggregates sectors G to U.

The circularity material use rate (CMUR) is also used an additional indicator for the waste subdimension. The CMUR represents the proportion of material resources used in an economy that come from recycled or reused materials. It is worth noting that circularity can be considered as a transversal dimension to the input/process/output system, but the CMUR was included under the process dimension due to its link with waste collection.

Indicators for the emission subdimension rely on Air emission data. This data can be retrieved from Air Emission Accounts, which measure air emissions in physical terms (e.g., tonnes) by NACE sectors following the resident approach. Air Emission Accounts provide data on a wide range of air pollutants, such as greenhouse gases (GHG) and particulate matter (PM10, PM2.5). These air pollutants are aggregated by Eurostat into three categories based on their negative effects on the environment. The categories are 'Global warming potential' (GHG), 'Acidifying gases' (ACG) and 'Tropospheric ozone precursors' (O3PR). The three categories together with the series on particulate matter smaller than 2.5 micron are used to create four indicators expressed as sectoral VA per one tonne of emission (the inverse of emission intensities).

In the occupation pillar, the task subdimension relies on task data from the Occupational Information Network (O*NET) database, which provides detailed information on occupations across the U.S. economy. O*NET is the reference source of information on tasks, and as such, it is a well-known dataset (Vona, 2021; Bachelot, 2024a). O*NET collects data on tasks (i.e. job specific actions and duties) divided into core and supplemental tasks. In general, each occupation is assigned between 10 and 25 tasks.

The occupational classification used in O*NET builds on the Standard Occupation Classification (SOC) of the Bureau of Labor Statistics and it is therefore required to use a crosswalk to match the

SOC to the ISCO classification used in Europe. I use the crosswalk provided by O*NET, which requires additional data adjustments discussed in Annex B1.4.

Following these adjustments, an indicator for green tasks is obtained as usually done in the literature, by computing the share of green tasks among the total number of tasks by occupation.

Indicators for the process dimensions are displayed in Table 2 and discussed in more detail in Annex B1.3 and B1.4.

Table 2: Process indicators

Subdomain	Source	Series name	Series code	brief description	indicators	remarks
waste-circularity	Eurostat	Generation of waste by waste category, hazardousness and NACE Rev. 2 activity	env_wasgen	Data on the generation of waste is collectedat national level, as per EU regulation. Waste generation is expressed in tonnes and is classified in sources (NACE 1-digit) and categories (European Waste Classification).	1) VA per 1 tonne of waste 2) share of chemical waste by sector 3) share of recyclable waste by sector 4) share of hazardous waste	service sectors aggregated as G-U
waste-circularity	Eurostat	Circular material use rate	env_ac_cur	The CMUR is an indicator of materials circularity in the economy and is computed as the share of materials obtained from recycled wastes over the total amount of material used.	1) CMUR by MS	MS level
emissions	Eurostat	Air emissions accounts by NACE Rev. 2 activity	env_ac_ainah_r2	This data set reports the emissions of greenhouse gases and air pollutants by sectors. Air pollutant are aggregated depending on their environmental pressure. The 3 categories are acidyfying gases (ACG), greenhous gas (GHG) and Tropospheric ozone precursors (O3PR). Data on emitted particulate matters smaller than 2.5 micrometer is also available (PM2_5).	1) VA per 1 tonne of ACG emission 2) VA per 1 tonne of GHG emission 3) VA per 1 tonne of O3PR emission 4) VA per 1 tonne of PM2.5 emission	-
Tasks	O*Net	-	-	The O*Net database provides information on the task contents of occupations. A green tasks classification is provided.	1) share of green task by occupation	-

Source: Own elaboration based on available information and the relevant literature.
Note: See Annex B1.3 and B1.4 for additional details on the datasets and indicators. The list of 1-digit NACE code can be accessed from https://en.wikipedia.org/wiki/NACE_rev2.

The final set of indicators for the green production subdimension are obtained from three different data sources. First, from the well-known environmental goods and services sector (EGSS) dataset, provided by Eurostat. This dataset allows for measuring the contribution of the green economy to the overall economy. Green economic output consists in activities related to environmental protection, aimed at preventing or mitigating pollution and conserving ecosystems, and resource management activities concerned with the sustainable use and management of natural resources. One indicator is created based on the total VA from green activities by NACE 1-digit sectors and expressed as a share of VA. Information for sectors G, H, I, K, L, N, Q, R, S, T, U is usually aggregated.

The remaining two indicators are constructed from data sources already used to generate indicators for the input dimension. These are the eco-innovation index (EC: Directorate-General for Research and Innovation, 2024) which includes an eco-innovation output dimension. This dimension is meant to capture the results of eco-innovation activities, and reflect the extent to which green innovations translate into practical applications, technologies, and market success.

The last indicator is constructed from the Physical Energy Flow Account already used for the energy subdimension. Instead of relying on the 'use' category, I rely on the 'supply' category to compute the share of non-green energy products among the total supply of energy products.

Data for all indicators is collected starting from 2006 up to the most recent year available (usually 2022). Many indicators are not available this far in time, which generates missing values. I present how to deal with this issue when discussing the methodology in Section 2.3. Data for LU and MT is missing for many indicators and these two MS are therefore dropped from the sample. NACE sectors T - Activities of Households as Employers, and U - Activities of Extraterritorial Organisations, are also not included in the analysis.

Table 3: Output indicators

Subdomain	Source	Series name	Series code	brief description	indicators	remarks
green production	Eurostat	EGSS	env_ac_egss2	The environmental goods and services sector (EGSS) accounts provide statistics on sectors creating environmental products understood as 'goods and services produced for environmental protection or resource management'.	(1) EGSS output as a share of sectoral VA	Data for sectors G, H, I, L, N, Q, R, S, T and U is aggregated into one broad sector (with values usually equal to 0) missing for: DE – O; SE – F;
green production	Eurostat	Energy supply and use by NACE Rev. 2 activity	env_ac_pefasu	Data from Physical Energy Flow Accounts (PEFA).	1) share of non-green energy products	MS level

Subdomain	Source	Series name	Series code	brief description	indicators	remarks
green production	EEA	Eco-innovation index	SUSO001	Eco-innovation refers to any innovation that reduces impacts on the environment, increases resilience to environmental pressures or uses natural resources more efficiently. The eco-innovation index is based on the eco-innovation scoreboard, which has indicators in five thematic areas	1) Index on eco-innovation output	MS level

Source: Own elaboration based on available information and the relevant literature.
Note: See Annex B1.5 for additional details on the datasets and indicators. The list of 1-digit NACE code can be accessed from https://en.wikipedia.org/wiki/NACE_rev2.

2.3. Quantitative Methodology

The quantitative methodology presented in this section covers steps 3, 5 and 6 of Nardo et al. (2005)[13], which address the weighting, aggregation and normalisation of indicators to create indices for subdomains, domains and the final composite indicator for green jobs.

For ease of exposition, it can be useful to consider weighting and aggregation taking place at two different stages of the composite indicator construction (Greco et al. 2019). First, indicators are weighted and aggregated to generate index values at the subdomain level. Second, weights are required at higher levels to aggregate subdomains together, followed by domains until the final GJI is obtained. These two different stages are addressed separately in Sections 2.3.1 and 2.3.3 .

This section starts with the presentation of the weighting scheme used to aggregate indicators and create subdomain indices. Different approaches can be used for weighting indicators (Nardo et al. 2005; Greco et al. 2019) but 'data driven' methodologies (e.g. principal component analysis or PCA) are particularly popular. PCA is a data reduction method, which summarises the variability in the dataset in few components. These (unobserved) components are constructed to be orthogonal to each other and ranked based on the share of the variance they contribute to explain. In general, the first component is used to create an index from the loadings associated to each observation. Provided that the indicators are correlated, PCA is an effective way to extract factors that drive common variations among the indicators. However, results from PCA should be considered with care as they can be

[13] Step 4 consists in a statistical multivariate analysis to better understand the association between all the indicators used for the construction of the composite index. This step is partially addressed in Annex B1, which shows the existence of a trend common to most series, consistent with the idea that economies are greening. Consideration related to this step are discussed further in Annex B3.

complicated to interpret and are not always proper for the construction of an index (Greco et al. 2019; Greenacre et al. 2022).

PCA is used to create the skill index for the labour subdomain. Since only one indicator is available for tasks, no weighting is required and only normalisation (discussed at a later point in this section) is relevant for the task subdimension. PCA can also be used for indicators included under the sector pillar but these series contain a time dimension (contrary to occupation data), which allows for exploiting variations between MS and sectors. As a result, subdimensions indices in the sector pillar are created using dynamic factor models (DFMs).

DFMs (Doz et al. 2012; Bańbura & Modugno, 2014; Poncela et al. 2021) are related to PCA[14] and are commonly found in (short-term) macroeconomic analysis where there are used to forecast and create indicators of the business cycle (Stock, 1989; Mariano & Murasawa, 2003). DFMs extract unobserved components common to a large number of series. Given that the interactions of indicators with the MS and sector dimensions results in more than 1000 series for some subdomains (e.g. energy, air emissions), DFM appear as a relevant framework to construct indices at the subdomain level through the extraction of the component driving common variations in the selected indicators.

Furthermore, DFM present the advantage of allowing performance of step 3 and 5 jointly rather than separately. These models are usually estimated using the Expectation Maximisation (EM) algorithm (Shumway & Stoffer, 1982). The EM algorithm relies on the Kalman filter and smoother which provides a straightforward solution to account for missing values. Diagnostics checks on the presence of outliers can also be obtained through the computation of smoothed residuals (Durbin & Koopman, 2012). Outliers can subsequently be corrected by including indicator variables (i.e. additive outliers, Chen & Liu, 1993) in the specification of the DFM. The EM algorithm is also very useful to estimate models with large number of parameters, though it is also known to converge rather slowly in the vicinity of the optimal solution. The algorithm guarantees convergence to a local maximum, which implies that initialisation is a crucial step for parameter estimation.

Moreover, the DFM specification can be extended to include explanatory variables. It is further possible to impose (linear) constraints on the parameters (loadings), which can be useful when the

---

[14] In fact, under a static factor model and other restrictions discussed in e.g.Bai & Ng (2013), results obtained from the factor model will be identical to PCA.

number of parameters to estimate is large. The model can also be used to forecast indicators and to provide values for normalising indicators (Section 2.3.2). Whether for PCA (labour index) or the DFMs, indicators are standardised to have zero mean and unit variance.

The remainder of this section presents the DFM (Section 2.3.1), explains how the estimation results are used to create the subdomain indices (Section 2.3.2), and concludes with the procedure to aggregate subdomains and domains into the final green job composite indicator (Section 2.3.3).

2.3.1 Dynamic factor model

The DFM is specified with only one factor assumed to evolve according to a first order autoregressive process[15]. This simplifies the exposition and makes the model more parsimonious since it should not be forgotten than the time window for the computation of the index (2006-2023) is relatively short in comparison to the number of periods used in standard application of the DFM. See Bańbura & Modugno, (2014) and for a complete presentation of DFM.

Let $Y_{i,j,t}$ denote the value of indicator Y (e.g. energy productivity, land use) in MS i and sector j at time t, and $y_{i,j,t}$ its standardised value. The DFM is then written as follows:

$$y_t = \Lambda f_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, H) \tag{1}$$

$$f_t = \mu_t \beta + \rho f_{t-1} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \tag{2}$$

The $n \times 1$ vector of standardised indicators at time t, y_t, is assumed to be linearly related to the unobserved factor f_t through the vector of loadings Λ. In addition, the factor follows an autoregressive process of order 1 and the specification allows for the possibility to include exogenous variables in the vector μ_t.

The model can be written in state space form (Durbin & Koopman, 2012) and the parameters (contained in) Λ, H, β, ρ and σ_η^2 can be estimated using the EM algorithm. The Kalman filter and smoother are used to evaluate the log-likelihood and perform the E-step of the EM algorithm The verification for outliers can be performed after estimation through the computation of smoothed residuals with an absolute value greater than a pre-defined critical value (Chen & Liu, 1993)[16].

---

[15] The specification does not include dynamic in the observation equation as well.

[16] The outlier detection and correction procedure will be performed for a future version of the composite indicator.

A particular concern in the current set-up relates to the fact that the number of variables is much greater than the number of time periods (i.e. $n > T$). Through a simulation exercise, Poncela et al. (2021) show that when n is large relative to T, the estimated factors and its associated forecast error variance obtained through the Kalman filter and smoother cannot be relied upon. The main issue appears to be related to the forecast error variance, which approaches zero as n increases. However, the forecast error variance obtained from one run of filter and smoother does not account for the uncertainty associated with parameter estimation. Given that the number of parameters to estimate increases with n, one could anticipate a substantial uncertainty associated with point estimates in the current set-up with $n > T$.

Therefore, the simulation results reported by Poncela et al. (2021) suggest that obtaining standard errors for the estimated parameters is very important in the case when $n > T$. These estimates could then be used to simulate the factor distribution by drawing from the parameter distribution, running the filter and smoother and extracting the factor, repeating this process for each draw. However, there are no straightforward solution to obtain standard errors when the EM algorithm is used (Poncela et al. 2021) and the evaluation of the uncertainty associated with the extracted factor is left for a future version of the index[17].

An important takeaway from the visual analysis of the indicator series displayed in Annex B2, is that many series across MS and sectors are characterised by a common trend consistent with the idea that economies are greening (e.g. higher use of renewable natural resources and green energy products as inputs, increase in the issuance of green bonds, increase in VA per one tonne of emission). It is because this trend is observed across a wide range of (green) indicators, and is common to many MS and sectors, that it is interpreted as a greening factor driving the co-movement of indicators[18].

The DFMs are then used to extract this factor and construct indices. Standardised series are included in levels (not in first-differences) and a linear time trend is added to the DFM specification in equation

[17] Note that a solution to this issue could be to numerically evaluate the hessian matrix at the EM solution using an optimisation routine available in many statistical software. The inverse of the hessian, known as Fisher's information matrix, can then be used as an estimate for the standard errors. It remains to be seen whether this evaluation is feasible when the number of parameters is large.

[18] The fact that indicators in MS and sectors might co-move heterogeneously with this green latent factor will be reflected in the estimated loadings.

(2). This is a rather naïve specification to account for the non-stationarity present in the series and more developed approaches could be employed. As will be seen in Section 2.4, the estimated factors do generally show an increasing trend suggesting that the DFM extracts the green component that seems to be driving the evolution of indicator series.

Behind the interpretation of the factor also lies the question of identification. It is well-known that the DFM is identified up to a rotation under the specification defined in equations (1)-(2) (Bańbura & Modugno, 2014; Poncela et al. 2021). Additional r^2 restrictions (with r the number of factors) need to be imposed for full identification of the factors and loadings. This should amount to one restriction in our set-up but the inclusion of the linear time trend also requires a an additional restriction (Bai & Ng, 2013). It is not very difficult to apply such (ad-hoc) restrictions in the context of the EM algorithm, but it is important to note that full identification is actually not required for the computation of subdimensions indices as these will rely on the quantities define by the product Λf. As a result, no restrictions were applied for the current version of index, which should be remembered when interpreting the DFMs estimation results in Section 2.4.1.

2.3.2 Normalisation of the indicators and computation of subdimensions indices

For each MS and each sector, the estimation results from the DFM are used to generate index values for each subdimension of the sector pillar (e.g. energy, capital, land use). This is done using equation (1), the estimated loadings $\widehat{\Lambda}$, the smoothed factor, $\widehat{f_t}$, retrieved from the Kalman smoother:

$$\hat{y}_t = \widehat{\Lambda}\hat{f}_t \tag{3}$$

The predicted values for the indicator constitute the core quantities around which the subdomain indices are constructed. As explained in the previous section (and as will become apparent in Section 2.4), the smoothed factor $\widehat{f_t}$ captures the common trend present in all the series. Hence, a high value will be obtained for series that strongly comove with the green component. This can be desirable as it implies that MS and sectors having strongly improved their 'environmental performance' according to the indicator considered (positive trend), will be assigned a higher score.

However, relying only on the trend could also generate misleading results since the level (i.e. the mean) of the series also contains important information on the indicator. For example, there are cases in which a series for an indicator does not show any trend over the sample period leading to predicted values close to zero ($\hat{y}_{i,j,t} \approx 0$ for MS i and sector j), while the original level of the indicator $Y_{i,j,t}$ could actually be relatively high.

Therefore, the predicted values, $\hat{y}_t$ are transformed back using the original series' mean and standard deviations[19], collected in the vectors $\bar{m}$ and $\bar{\sigma}$:

$$\hat{Y}_t \quad = \quad \bar{m} + \hat{y}_t \odot \bar{\sigma} \tag{4}$$

Where the symbol '$\odot$' is used to denote the element wise (or Hadamard) product.

Before aggregating the different indicators $\hat{Y}_t$ to obtain the subdomain indices, the indicators are normalised to lie between 0 and 100. This accounts for the fact that the $\hat{Y}_t$ might take on very different values across indicators after being transformed back to their original levels (e.g. between energy productivity and the share of green products).

A natural approach to normalising these indicators would be to rely, for instance, on some of the targets set out in the Environment action programme to 2030. However, these targets are usually not available for all indicators and are rarely set at sectoral level such that it is unclear whether applying these uniformly across all sectors would constitute a fair and relevant approach.

As a result, I use the DFM to project the factor up to 2030 and use this forecast to construct indicators values for the forecasted years (i.e. 2023-2030) using equation (4). These forecasts are simple and reflect the linear time trend, but they can be used to obtain a minimum and a maximum values for normalising indicators:

$$\tilde{Y} \quad = \quad \left(\hat{Y} - \min\left(\hat{Y}\right)\right) / \ \max\left(\hat{Y} - \min\left(\hat{Y}\right)\right) \times 100 \tag{5}$$

The maximum of the indicator is obtained across the entire sample and usually corresponds to the 2030 forecast values given the positive trend in the factor. However, this approach is not really appropriate when the indicator distribution is highly skewed[20]. For instance, this is the case of air emissions and the normalisation based on the entire sample would lead to indicators close to 0 for

[19] Note that a potentially interesting extension to the current DFM specification could be to jointly estimate the means and the factor.

[20] The skewness present in some indicators' distributions could be (partially) addressed by transforming the data (e.g. log, square root). However, these transformations make it impossible to interpret the subdimension indices as done below.

sectors A and B, which account for the bulk of air emissions, and close to 100 for the remaining sectors (a similar argument holds for land use and sector A).

As a result, it was decided to compute the minimum and maximum at the sectoral level across MS and time. This approach also suffers from limitations as an indicator could obtain a value close to zero, reflecting its poor sectoral performance when compared to the same sector in other MS, while this sector's performance could still be good in absolute terms when compared to other sectors[21]. In general, this approach will tend to smooth out sectoral differences and this potential effect should be kept in mind when analysing index results in Section 2.4.

Nevertheless, it should be noted that this normalisation has the advantage of offering a simple way to interpret the normalised indicators as the percentage point difference between the indicator and the target value (i.e. usually the 2030 value of the best performing MS for the given sector). For example $\tilde{Y}_{i,j,t} = 50$ means that the indicator in MS i and sector j is half of the target value set for sector j.

The index for subdimension S is then obtained by averaging each relevant normalised indicator:

$$I_{i,j,t}^{S} = \frac{\sum_k \tilde{Y}_{i,j,t}^{k}}{k} \tag{6}$$

For example, the energy subdimension is made of $k = 5$ indicators (e.g. energy productivity, share of green products, … see Table 1) and the normalised predicted values for each of these indicators, $\tilde{Y}_t$, are averaged to obtain the energy subdimension index.

Using the average to compute subdimensions indices is simple and implies full compensation between indicators (see the discussion in Section 2.3.3 as well). It further gives a simple interpretation to the subdomain indices. Given that $\tilde{Y}_{i,j,t}$ corresponds to the percentage point difference between the indicator and the selected target values, the subdimension indices correspond to the average percentage point difference across the k indicators.

Two additional points are worth mentioning regarding the computation of $I_{i,j,t}^{S}$. First, when indicators are only available at MS level then the same predicted value $\tilde{Y}_{i,t}$ is used for all sectors j. Therefore, these indicators will tend to smooth differences across sectors and it will be interesting to perform

[21] An alternative could be to normalise indicators based on both approaches (i.e. full sample and sector sample) and take the average of the indicators. This solution will be explored in a future version of the index.

some robustness checks on these variables. Similarly, when data on the service sector is available as an aggregate sector (i.e. usually sector G to S) then the predicted value obtained for this aggregate group is assumed to apply to all individual sectors.

Second, when an indicator is missing for a specific MS and sector (See Table 1, Table 2 and Table 3)[22], the average is computed based only on the available indicators. This implies that some subdomain indices for certain MS and sectors are computed from a different number of indicators. This is a concern that only affects 65 series out of the 5000 included in the analysis. This represent slightly more than 1% of the series and it should be noted that most missing series can be linked to the land use indicator.

For the subdomains falling under the occupation pillar, namely the labour and task subdomains, it should be remembered that the indicators are log-transformed. The three skills indicators are aggregated by computing the score obtained from PCA, whilst the task subdimension only include a single task indicator. The subdomain indices are then normalised using the maximum and minimum observed across the sample of occupations.

Given that these indicators are log-transformed, these subdomain do not have the same simple interpretation that characterises the subdomains falling under the production (sector) pillar. However, they will be increase when the original indicators do so.

2.3.3 Aggregation at the domain level and higher

Equipped with indices for each of the subdimension, I can then generate indices for the dimensions, namely, input, process and output for the sector and occupation pillar respectively, followed by the aggregation of these three dimensions to obtain the sector and occupation indices. Finally, the sector and occupation pillar indices are aggregated together to obtain the GJI.

This step should be guided by a number of considerations related to the conceptual framework and the nature of the indicators. There exists various approach and an important aspect relates to whether aggregation should be compensatory or not (Nardo et al. 2005; Greco et al. 2019).

For this first version of the index, I follow a simple approach and compute all domain as well as the sector, occupation and final GJI by averaging the relevant components. This assumes perfect

[22] Note that a series is considered to be missing if less than four observations are available. See Section 2.4.

substitution between the components making the different indices (e.g. a decrease in one point in the energy index can be compensated by a 1 point increase in the capital index).

Moreover, using the arithmetic mean implies that components making the index are implicitly assigned equal weights. In the current set-up for instance, it might be relevant to assign a lower weight to the land dimension given the low number of available observations for the indicator on land use. Different weights could also be applied at higher levels to give more importance to certain dimensions (e.g. green output) if it is believed that such dimensions should be given a more prominent role to determine the greenness of a job. These weights could be determined after expert consultations, as done in the case of the gender equality index (European Institute for Gender Equality 2017). Alternatively, data driven methodologies (e.g. PCA, DFM) could also be used. Testing the sensitivity of the index values to different weighting and aggregation schemes constitute an important task left for future work.

Yet, despite its simplicity, the arithmetic mean offers the advantage of maintaining the interpretation of the index values under the sector pillar as the average percentage point difference between indicators and their target values. It also allows for simple decomposition of the final GEI in terms of different components (Section 3.2).

2.4. Results

This section presents the results obtained from applying the methodology described in the previous section.

2.4.1 sector pillar

Before briefly discussing estimation results, it is worth mentioning that the DFM cannot be applied to all series. In addition to series missing for certain sectors and MS, some series only contain very few observations, which can raise concern on the estimation of their associated loadings. As a result, series with less than four observations are not included in the estimation of the DFM and are considered to be missing[23]. Some series (e.g. the share of non-renewable energy in BE) can be constant, at values usually equal to zero. This is only a concern for estimation and these series can

[23] Land use only includes four observations maximum and this restriction is therefore loosened to any series with less than three observations available.

still be used for the computation of the subdomain indices by setting $\hat{Y}_{i,j,t} = \overline{m}$ with $\overline{m}$ generally equal to 0.

The DFM is estimated at the subdomain level by including all series falling under the relevant set of indicators (e.g. all series at MS and sector level for the five indicators making the energy subdimension). Given that the aim is to extract a single (greening) factor, it is possible to argue that estimation should take place at a higher level than the subdomain level (e.g. domain or even the sector pillar level). However, this would entail estimating close to 10 000 parameters and this approach was not pursued in this version of the index. The only exception concerns the emission subdimension for which the DFM estimation using the four indicators jointly (e.g. GHG, ACG) led to an estimated autoregressive parameter close to one. As a result, estimation under the emission subdimension was done at the indicator level (i.e. four different DFMs have been estimated for each of the four indicator). It is worth noting that non-stationarity is not a problem for the Kalman filter and the EM algorithm provided the filter is initialised properly (Poncela et al. 2021).

Regarding initialisation of the EM algorithm, I follow Bańbura & Modugno (2014) and use PCA to initialise the algorithm. This applies to all subdimensions except land use for which PCA initial values lead to improper results[24]. Therefore, I use different set of starting values for land use and retain results leading to an extracted factor displaying a trend. The issue with land use could be explained by the small number of available observations (maximum four) and allows for reiterating the need to consider this subdimension with care.

Finally, the convergence criteria is set following Bańbura & Modugno (2014) as:

$$\frac{\mathcal{L}(i) - \mathcal{L}(i-1)}{(|\mathcal{L}(i)| + |\mathcal{L}(i-1)|),} < \quad 1.0e^{-7}$$

Where $\mathcal{L}(i)$ is the value of the log-likelihood in iteration i evaluated using the Kalman filter.

<u>*DFM estimation results*</u>

Estimation results for each subdomain are displayed in Annex C1 and consists of three figures displaying the estimated factor, as well as the loadings averaged across MS and sectors respectively.

[24] The EM algorithm should be increasing at all iterations until convergence (Shumway & Stoffer, 1982), which is not the case when PCA is used to initialise the algorithm for land use.

The first point to note from these results is that extracted (green) factors are increasing through time for all subdimensions except land use[25]. This indicates that, despite the lack of identification, the DFMs appear to extract the common trend characterising the selected indicators, and its sign is consistent with the idea that sectors are greening.

The remaining results display average loadings by MS and sectors, and are generally in line with the descriptive evidence discussed in Annex B1 and displayed in Annex B2. Positive average loadings usually signal that the indicator has been increasing while the opposite holds for negative loadings. The higher is the loading, the stronger is the co-movement between the series and the green factor.

A deeper analysis of these results, going beyond average loadings, would likely reveal some interesting facts on the greening process taking place in different sectors and MS. However, given the lack of identification and the fact that the main use for the DFMs output is to construct subdimension indices, these results have not been analysed further.

Subdomain, domain and sector indices

Figure 3 displays average indices for the input dimension by sector and MS. From this figure, it is worth first mentioning that there is more variations in the means across MS than sectors. This observations holds for other components of the sector pillar (Figure 4 and Figure 5), and is likely reflecting some of the methodological choices made in Sections 2.2 and 2.3. In particular, the choices related to the inclusion of MS level variables and to the normalisation of indices. The normalisation decision is important as it implies that the various indices corresponds to average[26] percentage point differences between indicators and their target values set at sectoral level. As a result, the similar average index values across sectors indicate that all sectors are at a similar distance (in terms of percentage point) from their targets across the range of indicators considered.

Nonetheless, the sectoral rankings for each (sub)dimension tend to be aligned with the expectation that industrial sectors (including agriculture and construction – A to F) would generally obtain lower index values compared to the service sectors. In panel a), average indices lie between 26 and 39 with

---

[25] For land use, the factor is decreasing but it should be noted that MS and sectors which saw an increase in the indicator (decrease land use) are associated with negative loadings (Figure 56

[26] Actually, indices for the sector pillar and at the domain level are weighted averages, with weights reflecting the different number of indicators in each subdimension.

sectors C and F obtaining among the lowest average values. On the other hand, sectors A, B and E feature among the best performers due to high score obtained from specific indicators[27]. At MS level (panel b) in Figure 3), average index values range between 20 (RO) and 57 (SE). These results can again be traced back to certain indicators (share of green products for SE, share of non-green products and share of non-renewable energy for RO).

Regarding the capital subdimension, it was already pointed in Section 2.2 that this dimension only included one indicator available at sectoral level (investment in climate mitigation) different from zero for sectors C, D, F, H, J and M[28]. This explains why these sectors obtain the largest average values in panel c). In panel d), SE is ranked first again with an index value of 51 followed by DK and FI. In general, the ranking appears to be correlated with MS' wealth (i.e. GDP per capita), much more than in panel b), reflecting perhaps the more developed capital markets in these MS.

For the land subdimension, panels e) and f) show that average scores are much higher than those obtained for the energy and capital subdimensions. This reflect the skewness of the land use distribution, with certain sectors in specific MS occupying a larger share of land, and influencing the normalisation of the indicator (Section 2.3.2). This is typically the case of mining which accounts for less than 0.5% of land use in the vast majority of MS, but represent 2.3% of land use in IE and 0.9% in EE. As a result, sector B attains high scores in most MS and low ones in IE and EE. It is further important to remember that this subdimension is constructed on a single indicator informing on the share of land use, without any reference to environmental consequences stemming from the activity.

It is nonetheless interesting to point form panel e) that sector A, which represents more than 75% of land use on average, obtains the lowest index value[29]. At MS level, NL is ranked last, though this number should be considered with care given that it is seems driven by the increase in land use

---

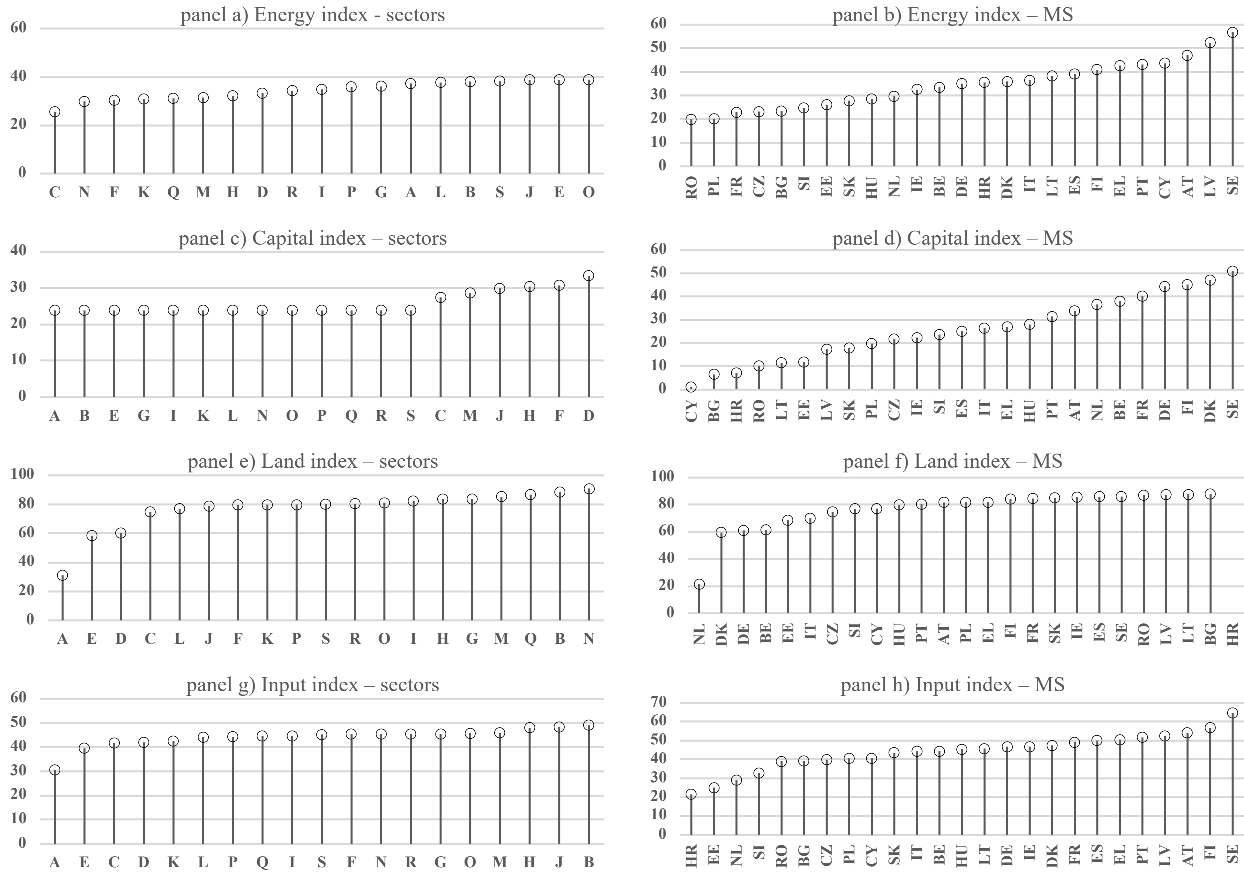
[27] At the indicator level, $\hat{y}_t$, sector E is the best sector on average in terms of the share of green energy products used. Sector A performs well in terms of green energy products as well, but is the best performer, with sector B, for the share of non-green energy products. These results are not reported in this document.

[28] Index values for the remaining sectors are different zero, which represents the contribution of the MS level variables (i.e. share of green bonds and eco-innovation input index).

[29] The normalisation of the index for sector A is affected by CY, which appears as an outlier with sector A only representing half of land use in this MS.

observed in 2018 (last data point available). BG and LT attain the highest rank with index values slightly below 88. Data for HR is not available.

Figure 3: Average indices – input dimension



Source: Own elaboration.
Note: https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Regarding the process dimension in Figure 4, results for the waste subdimension are such that low index values are found for industrial sectors (A to F). For service sectors, waste indicators are only available for sectors G to S aggregated, and this unique value is used for all the individual sectors. The waste index at MS level ranges between 38 in EL and 62 in FR with BE and NL also obtaining values above 60. It is interesting to note that some of the MS which ranked high based on the input index (e.g. FI, LV, SE) tend to rank lower on the waste index.

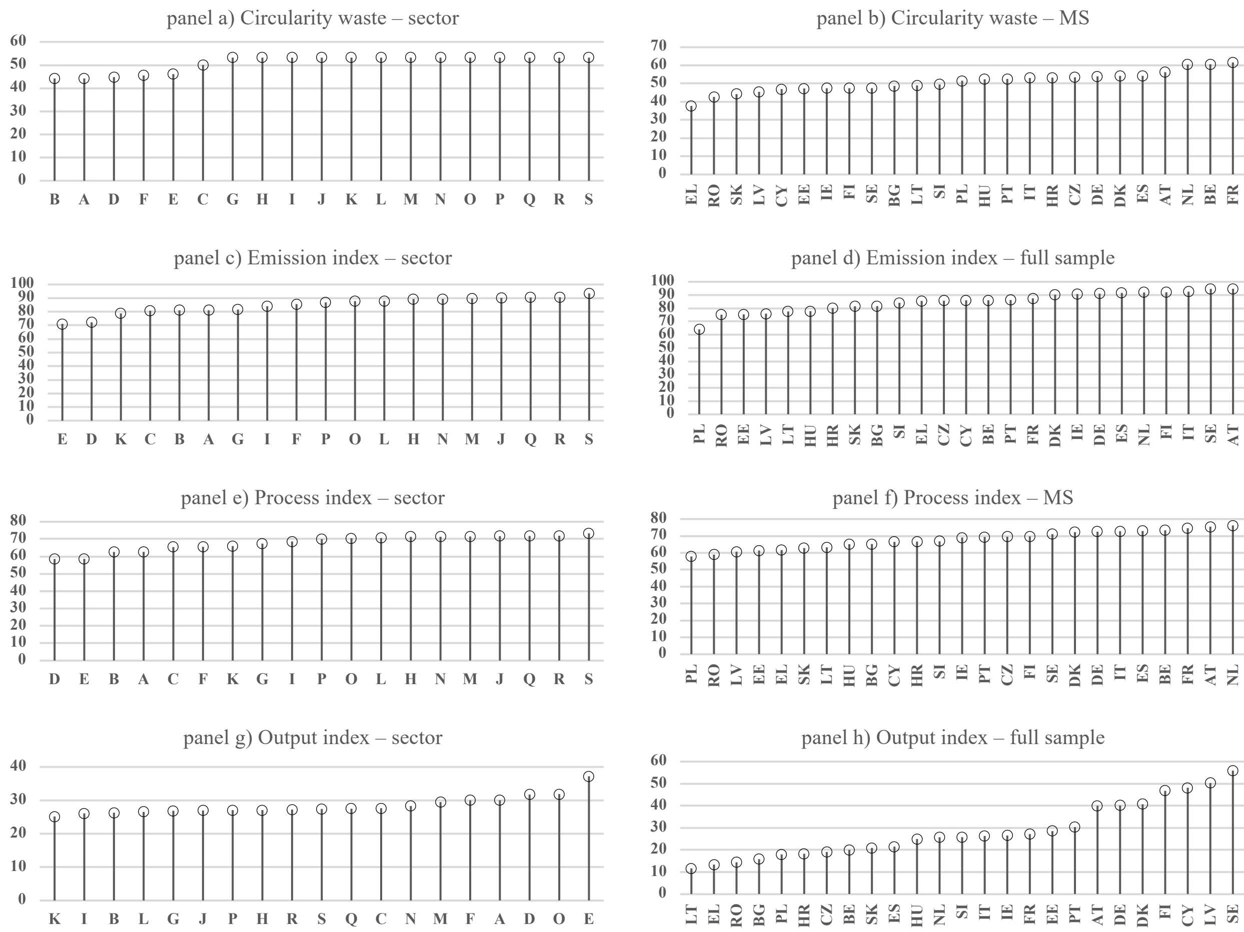
As is the case for the land use indicator, the distribution for emissions indicators remains highly skewed even when normalising indicators through sectoral values. This explains the high averages score in panels c) and d) (above 60) obtained for this subdimension. It also means that on average, emissions indicators are closer to their target values than it's the case for other indicators. In terms of ranking, industrial sectors, with the exception of sector K, are ranked the lowest, while service sectors obtain the best scores. With an average value of 71, sector E obtains the smallest value, while sector S is first with an average value of 93. By MS, countries appears to be ranked geographically with Western (including Southern and Northern) European MS obtaining higher index values than Eastern MS. With a value of 64, PL is ranked last, more than 10 points below RO in the second-to-last position (index value of 75). AT and SE obtain the highest scores for the emission index with average values close to 95.

The process index is computed as the average of the waste and emission indices and the results therefore reflect the points discussed above. In particular, industrial sectors obtain the lowest score between 60 (sector D) and 66 for sector F, whereas service sectors attain average values above 70, except for sectors K, G and I. With regards to MS, NL and AT obtain the highest index values as they perform well in terms of both waste and emissions. They are followed by FR and BE, which rank high on the waste index. PL is ranked last due to its very low score on the emission index, whilst RO and LV perform poorly relatively to other MS in terms of both waste and emissions.

The bottom panels in Figure 4 display results for the output dimension. The output dimension only features one subdimension (green production, see Figure 2) constructed from three indicators. Panel g) shows that index values at sector level range between 25 (K) and 37 (E). The ranking of sectors is different from the ranking obtained using the input and process indices as industrial sectors generally attain higher values than service sectors. This fact reflects that the EGSS VA tend to be greater in industry. The MS ranking is however more in line with what has been discussed until this point. SE, LV and FI obtain high average scores, while index values tend to be lower for Eastern European MS. It is interesting to point the significant break in index values between the top performing MS with an index above 40 (i.e. SE, LV, CY, FI, DK, DE and AT), and the remaining MS with average values of 30 or below. This seven MS tend to perform well in terms of the three indicators used to construct the output (sub)dimension.

The input, process and output indices are then used to construct the index for the sector pillar. Given that the arithmetic mean is used, the results are also in line with the points highlighted above. At sector level, the results show that despite their good performance on the output index, industrial sectors can be considered as the least green sectors. Sector A is ranked last with an average value of 41. Despite coming first on the input index, sector B is still classified among the least green sectors due its results for the process and output dimensions. Sector K, which poorly performs in terms of all three dimensions, also features among the lowest ranked sectors. Service sectors O, J and M appear as the greenest sectors with index values above 49.

Figure 4: Average indices – process and output dimensions



Source: Own elaboration.
Note: https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

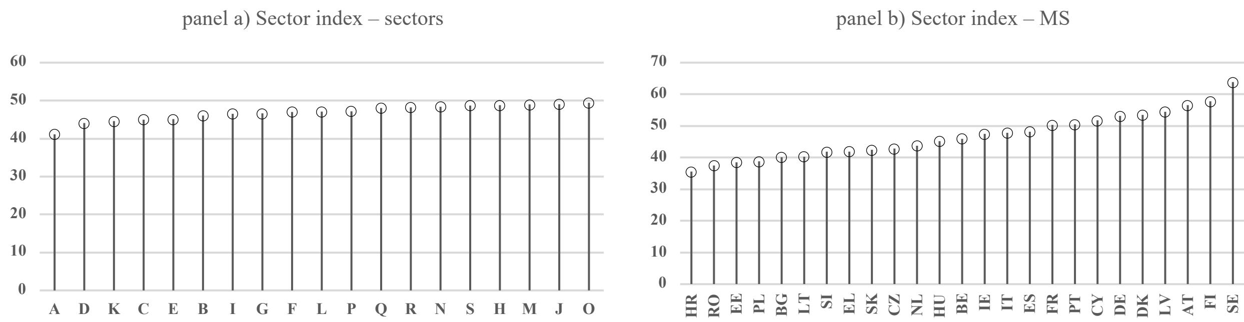
At MS level, SE appears as the MS with the greenest sectors on average. With a value close to 64, SE outranks FI by more than 6 points. AT and FI also performs well in terms of all three dimensions and it is therefore not surprising to see these MS obtaining high sector index values. LV would rank higher if it was not for its performance on the process dimension.

Note that HR (and NL) is probably penalised by their performance in terms of the land use indicators, which drives the low input index value discussed previously. The ranking for these MS should therefore be taken with a grain of salt.

Finally, Figure 6 and Figure 7 display the evolution through time of the sector index by sectors and MS. These figures are interesting as they reveal the greening taking place at sectoral level in most

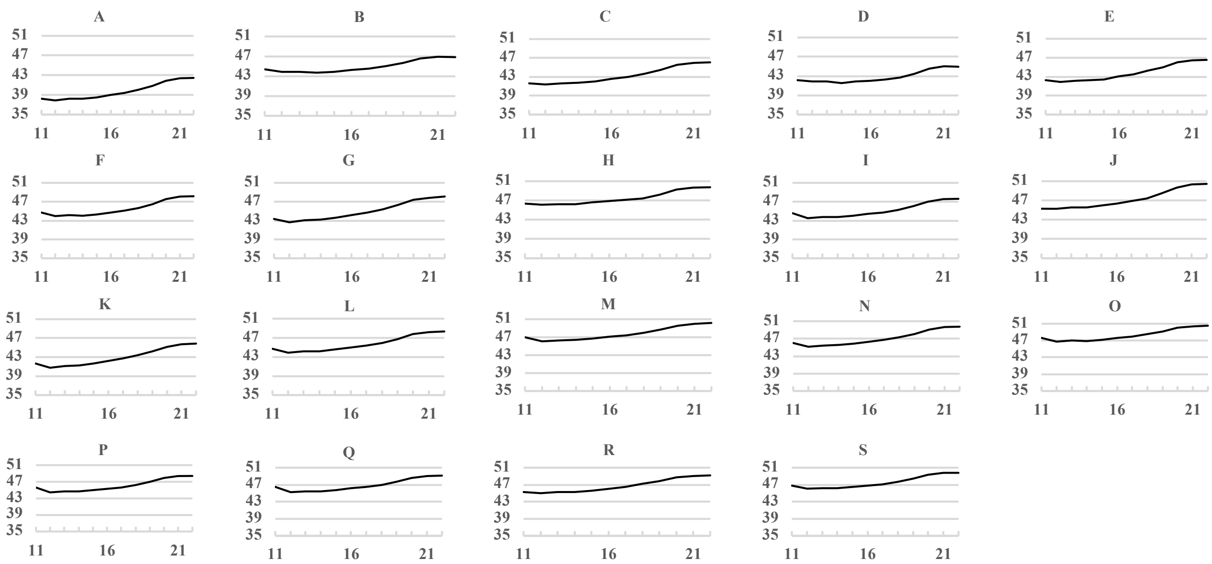
MS. Between 2011 and 2022, the highest progress has been registered for sector J (+5 point). Interestingly, sectors C, E, A and K are among the sectors showing the largest progress (between 4.1 and 4.5 points) suggesting a (slow) catch-up of the least green sectors over the period. The lowest increased in the sector index is observed for sector B (+2.5 points).

Figure 5: Average sector index



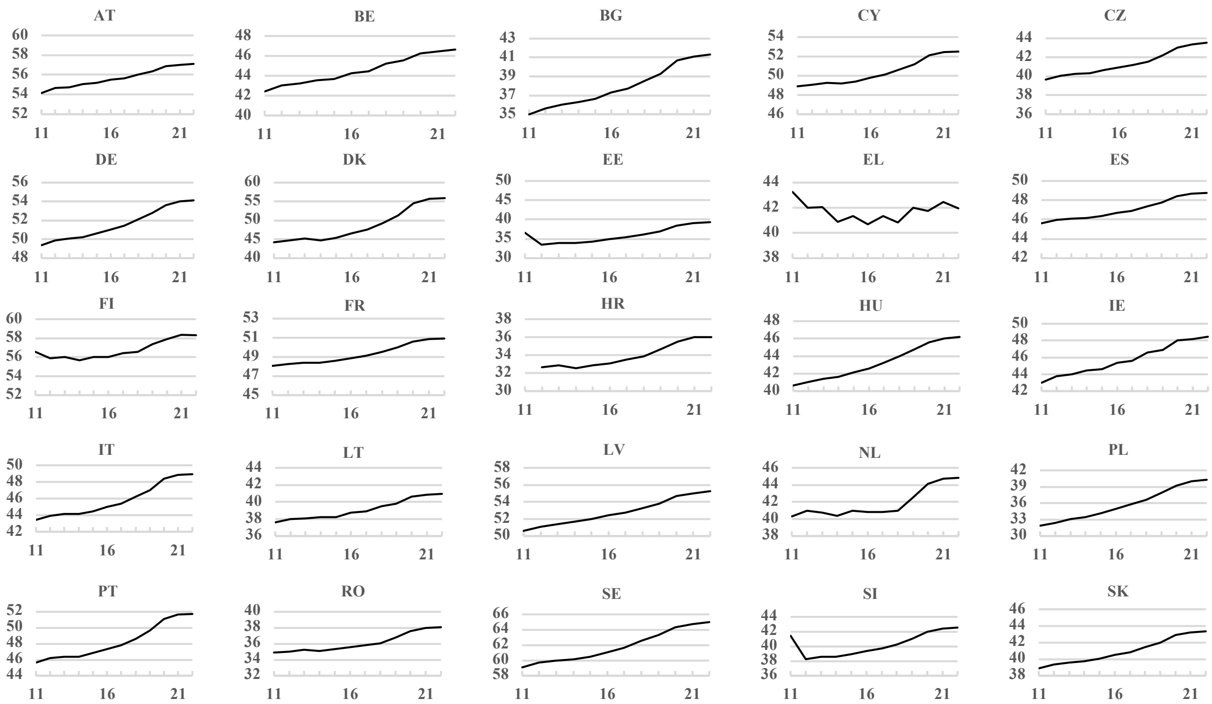
Source: Own elaboration.
Note: https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 6: Sector index by sector



Source: Own elaboration.
Note: https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 7: Sector index by MS



Source: Own elaboration.
Note: https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

2.4.2 *Occupation pillar*

Regarding the occupation pillar, it is first worth mentioning that BG and SI provide occupational data at the 2-digit level and the indices in the occupation pillar are therefore averaged at this level for these two MS.

In general, results obtained from indicators under this pillar are well-known. Moreover, the task indicator constitutes the only variable under the task and process (sub)dimensions and these index can directly be obtained as a rescaled version of the results displayed in Figure 46.

As a result, this section only briefly discusses the index retrieved from skill/knowledge indicators.

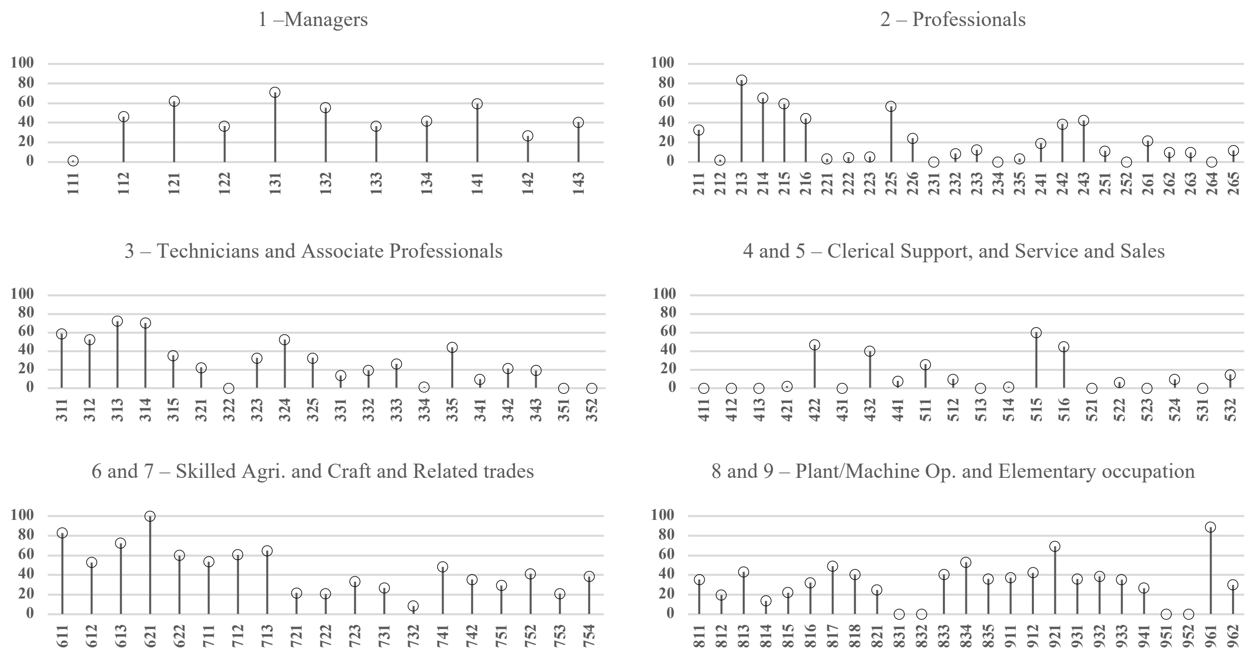
Skill/knowledge index

The three skills indicators are aggregated using PCA. Results can be found in Table 5 and do not appear to raise any cause for concern. The skill/knowledge index is then constructed as the score from the first component for each 3-digit occupation.

Results are displayed Figure 8. Considering that the skill indicators have been log-transformed before aggregation using PCA, the results tend to be similar to those reported by Maldonado et al. (2024) with, for instance, high values of the index for skilled agricultural occupations.

In the occupation pillar, the scores for the input and process dimensions are simply equal to the labour and task subdimensions. Index values for the occupation pillar are therefore obtained by averaging the green task and green skill/knowledge indices. Given the absence of variations across MS and times, the occupation index is discussed more extensively in Section 3.

Figure 8: Skill/knowledge index



Source: Own elaboration.
Note: See https://ilostat.ilo.org/classification-occupation/. for the list of occupation codes.

2.4.3 *Green job index*

The GJI is obtained as an average between the sector and occupation indices. The occupation index does not vary through time and across MS. As such, all the variations along these two dimensions will originate from the sector index already discussed in Section 2.4.1. Furthermore, in its current form before matching with EU-LFS data, the GJI includes index values for many irrelevant sector-occupation pairs. Hence, the gains from a deeper analysis of these results are limited at this stage and it is more interesting to directly move to the GEI results.

3. The Green Employment index

The Green employment index is obtained by merging the GJI to the EU-LFS using NACE rev2 and ISCO08 3-digit codes to the EU-LFS. The employment share for each sector-occupation pair can then be computed using sample weights provided in the dataset. The sample starts in 2011 when the ISCO-08 classification was introduced in the EU-LFS, and is further restricted to individuals aged 16-64 holding only one job. For each sector j and occupation l, the GEI, omitting the MS index i and time index t, can be computed as:

$$GEI = \sum_j \sum_l \omega_{j,l} GJI_{j,l} \tag{7}$$

Where $\omega_{i,j,t}$ is the employment share of the sector j and occupation l. Index for each (sub)dimensions can also be obtained using the same equation replacing the GJI with the relevant index.

Before commenting on the results, it is worth mentioning that given the work left in regards to robustness checks and sensitivity analysis of the index, the results discussed in Sections 3.1 and 3.2 should be seen as preliminary and considered with care.

3.1. GEI Results

Table 4 presents average index values over the 2011-2022 time period for each MS. According to the GEI, employment is the greenest in SE with a value of 45, followed by FI, AT and LV with respective values of 42 and 41. Most MS reach average values between 30 and 40, and employment is the least green in HR, EE, BG and PL, with values of the GEI below 30.

Before analysing further these results, it is worth reflecting briefly on the size of the GEI. The occupation pillar is constructed from (log-transformed) indicators, which increase with the shares of green skills, knowledges and tasks. The maximum contribution from this pillar would then be obtained if the skills and tasks indicators were set at their maximum sample values for all occupations[30]. As such, a value of 100 should not be seen as the desirable outcome, although fluctuations in this index remain relevant to judge how employment is greening. The rather small index values for the occupation pillar (between 15 and 36 in Table 4) highlight that many jobs do not feature any green skills or tasks, and when they do, the shares are small in many instances.

[30] Or at least some occupations representing 100% of employment.

On the other hand, indices obtained under the sector pillar maintain their interpretation as the (weighted) average percentage point difference between indicators and their target values set at sectoral values. Hence the value of 68 for SE implies that sectoral indicators for this MS are at a level representing two-thirds of the target indicators on average. In general, high values of the sector index signals that the MS performs well in terms of the indicators falling under the given (sub)dimension.

Table 4: Average Green Employment Index – 2011-2022

	GEI	sector									Occupation		
			input				process			Output		input	process
				Ener.	K	lu		Waste	Emission			L	task
AT	41	59	56	49	34	84	77	62	92	44	22	27	17
BE	33	45	41	24	40	58	77	65	88	16	22	26	17
BG	29	34	38	17	8	88	57	51	62	8	23	30	17
CY	37	53	39	39	1	77	65	52	78	54	20	25	16
CZ	32	40	36	15	22	72	68	58	79	15	23	29	18
DE	39	56	44	25	44	64	74	59	90	48	22	27	18
DK	35	50	43	25	49	55	72	61	82	34	20	24	17
EE	29	32	25	19	15	31	49	44	54	22	26	31	20
EL	31	39	48	38	25	82	61	46	77	7	22	28	16
ES	34	47	51	36	26	91	73	59	87	18	22	28	15
FI	42	61	59	39	47	90	74	60	88	51	23	27	19
FR	37	52	49	19	42	87	79	71	88	27	23	28	18
HR	27	31	18	31	6	-	62	59	65	13	23	29	18
HU	33	42	43	22	29	79	59	58	60	24	23	29	17
IE	35	47	47	27	23	91	70	49	90	25	23	28	18
IT	32	44	40	28	26	68	73	55	90	18	21	26	15
LT	31	36	42	30	11	85	63	51	76	4	26	32	20
LV	41	56	55	58	19	87	55	47	63	60	25	32	19
NL	31	42	26	17	37	24	78	66	90	21	21	25	17
PL	29	32	36	11	21	77	48	56	40	11	27	32	21
PT	34	47	49	37	30	81	63	54	72	29	21	26	16
RO	31	32	33	15	11	71	54	46	62	8	30	36	23
SE	45	68	68	60	52	92	71	49	92	65	23	27	19
SI	32	40	32	25	24	36	64	55	72	23	23	29	18
SK	31	39	40	21	17	82	61	47	76	16	23	28	18

Source: Own elaboration.

Taking a closer look at the two pillars and their (sub)dimensions, it is possible to see that SE attains relatively high values for all dimensions under the sector pillar with the exception of waste. To a lower extent, this is also the case for AT and FI, which obtain relatively high index scores for all the three dimensions. For LV, its high GEI value is obtained with one of the lowest average index score for the process dimension. Similarly, MS with the least green employment in terms of the sector index perform poorly (or average at best) for all the (sub)dimensions.

For the sector index, it is further possible to note that DE, CY, FR and DK attain values above 50, while the GEI for ES, IE and PT lies slightly behind this value. At the input level, CY, ES, PT and EL obtain relatively high values, partly because of their good score on the energy subdimension. This is not case for DE, DK and FR for which the index on this subdimension is rather low. This low value is compensated by their good performances in terms of the capital subdimension.

FR further obtains the highest index value for the process dimension, driven by its high score on the waste subdimension. Several MS obtain values above 70 for the process dimension, including NL, AT, BE, DE, FI, ES, IT, DK, SE and IE. In terms of output, the results are in line with those discussed in Section 2.4.1 and displayed in Figure 4, as the seven MS highlighted there are those obtaining the highest index values. MS such as ES, FR, IE, PT and NL, which performed rather well on the input and/or process dimensions, tend to attain lower values on the output index. The output index for these MS lie below 30, which represents less than half of the score obtained by the two best performing MS (SE and LV with average values of 65 and 60).

With regards to the occupation pillar, it is important to remember that these indices do not vary across MS and time. Hence, variations between MS reported in Table 4 only reflect differences in the occupational composition of employment. With the exception of LV, it is possible to note that MS obtaining the highest values for the occupation index (> 25) are those obtaining low values for the GEI (i.e. RO, PL, EE, LT, and LV). As discussed further in Annex B1.2 and B1.4, this results can be understood from the fact that important green jobs, in terms of skills in particular, are associated with (skilled) agricultural occupations (Maldonado et al. 2024). Sector A, where these occupations are mainly located, represents a high share of employment in these MS, but it is also the sector obtaining the lowest index value for the sector pillar (Figure 5). These effects originating from the sectoral distribution of employment are explored further in Section 3.2.

Finally, it is worth noting the certain overlap between the labour and task indices with MS generally obtaining high or low values for both indices.

Going beyond individual MS results,

Figure 9 and Figure 69 (in the Annex) reveal a steady increase between 2011 and 2022 of the GEI for all MS but EL[31]. This indicates that employment has been greening across the EU, albeit at a heterogenous pace across MS.

Figure 9 shows that DK registered the largest increase in the GEI (+9.2 points), followed by PT, BG and PL, with improvements ranging from 5.3 to 6 points. Furthermore, MS such as HU, IT and EE, also feature among MS registering substantial improvements in their GEI (between +4 and +5 points). In Table 4, these MS usually obtain an average or below average scores on the GEI, with BG, EE and PL featuring among the MS with the least green employment. Therefore, the substantial increase in these MS' GEI between 2011 and 2022 suggests a certain convergence during this period as employment appears to green faster in these MS.

In terms of the two pillars, it is clear from panels b) and c) that this progress was largely driven by greening taking place at the sector level. The occupation index increased only in handful of MS with the largest increase registered for SE, DK and HU (+1.5, +0.9 and +0.8 points, respectively). On the other hand, some MS registered a decrease in their occupation indices, particularly in RO. As already mentioned, this evolution reflects the decrease in the share of employment in Sector A (Maldonado et al. 2024). One should not infer from these observations that absent this change in the occupational composition of employment, the increase in the GEI would have been greater since this change is taking place together with a reallocation of employment from sector A (one the least green in terms of the sector pillar) to other greener sectors. Hence, this example shows that the greening process is not a linear one and as economies develop, moving away from sectors such as agriculture, the greening of employment can be expected to be slow.

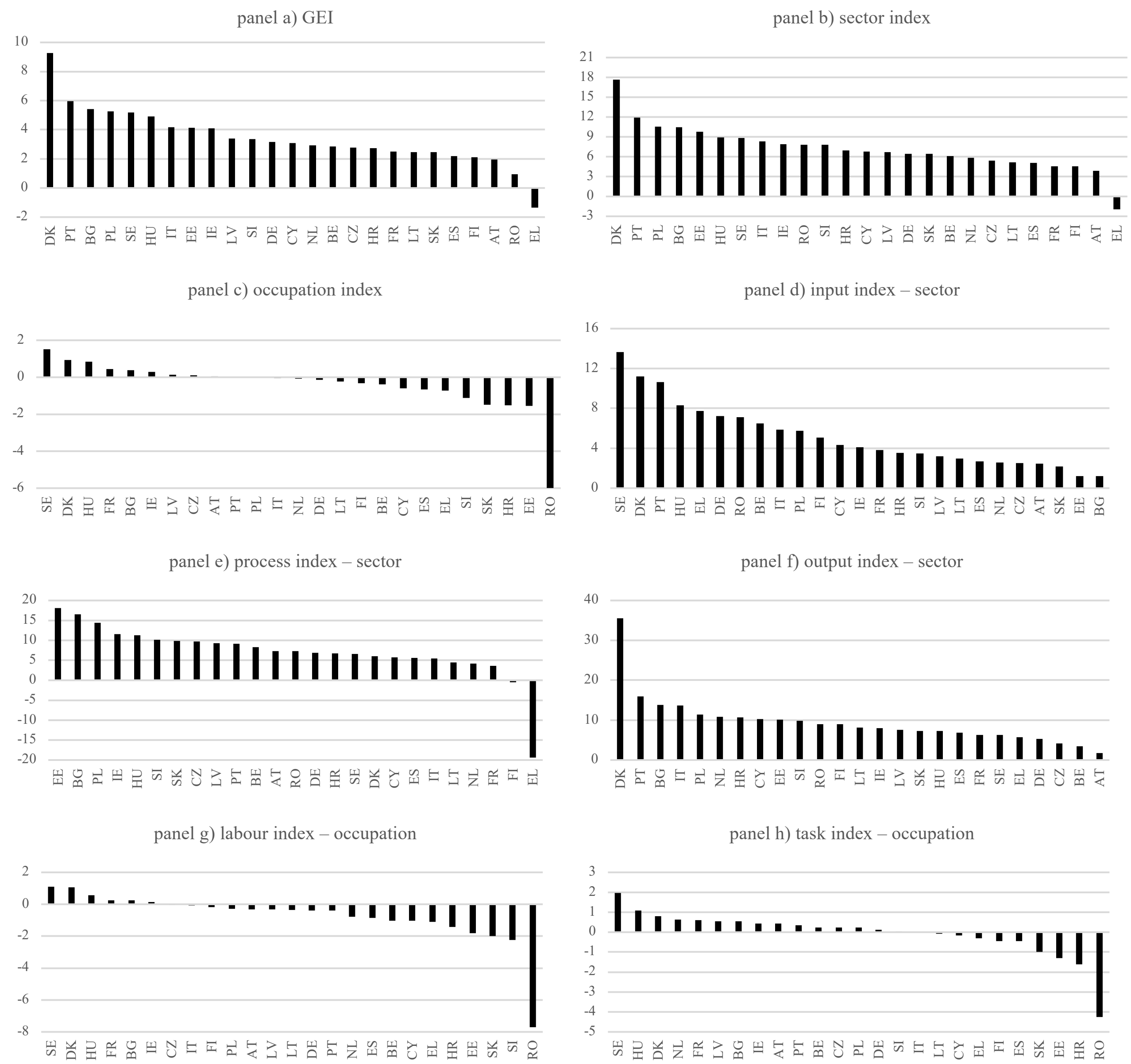
Finally, in regards to the sector pillar, it is worth noting that the increase in DK is largely driven by progress on the output and, to a lower extent, the input dimensions. This observation also holds for PT, which also registered a significant increase in its index value for the process dimension. This later dimension also appears as the main driver behind progress in the GEI for BG, EE and PL. The

[31] In the case EL, it is worth remembering that this decrease originates only from the waste dimension and its indicators (see panel e) for the process dimension in

Figure 9). It is not clear at this stage whether this indicates a genuine worsening of waste indicators or whether this trend can be explained by other reasons such as a break in data collection.

output indices also increased significantly for these MS, but it appears that progress made in terms of waste and/or air emissions is the main driver behind the greening of employment taking place in these MS.

Figure 9: Change in index values – 2011-2022



Source: Own elaboration.

3.2. Unpacking the GEI

This section presents a first attempt at better understanding the main factors driving the GEI. To do so, this section performs a small decomposition exercises focusing on cross country variations in the

GEI in three different sources, while the second proposes an analysis of sectors with the most substantial contributions to the GEI by MS.

3.2.1 *Decomposing cross MS variations in the GEI*

The variations in the GEI across MS can come from different origins. First, let us rewrite the GEI defined in equation (7) using the fact the GJI is obtained as the average of the sector and occupation indices:

$$\begin{aligned} GEI_i &= \sum_j \sum_l \omega_{i,j,l} GJI_{i,j,l} \\ &= \sum_j \sum_l \omega_{i,j,l} I_{i,j}^{Sector}/2 + \omega_{i,j,l} I_{i,l}^{Occup}/2 \\ &= \sum_j I_{i,j}^{Sector}/2 \sum_l \omega_{i,j,l} + \sum_l I_l^{Occup}/2 \sum_j \omega_{i,j,l} \end{aligned} \tag{8}$$

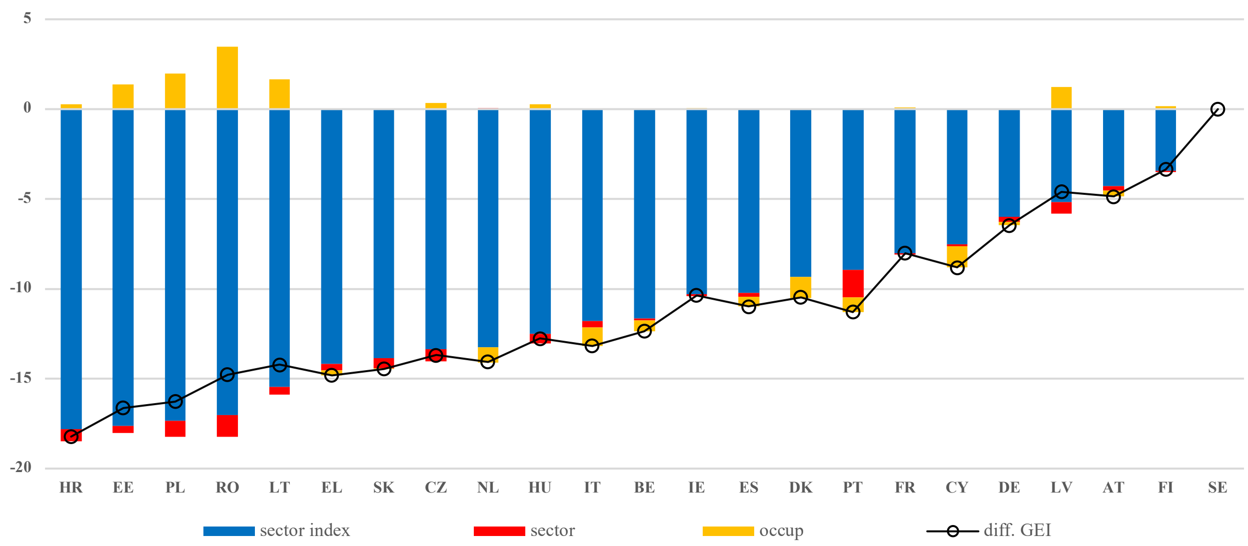
Taking one MS as reference (i.e. SE), the difference between the GEI in MS i and the reference MS can be decomposed into three terms:

$$\begin{aligned} \Delta^{SE} GEI_i = & \underbrace{\sum_j \left(I_{i,j}^{Sector} - I_{SE,j}^{Sector}\right)/2 \sum_l \omega_{i,j,l}}_{Sector\ index} \\ & + \underbrace{\sum_j I_{SE,j}^{Sector}/2 \sum_l \left(\omega_{i,j,l} - \omega_{SE,j,l}\right) +}_{Sectoral\ composition} \\ & + \underbrace{\sum_l I_l^{Occup}/2 \sum_j \left(\omega_{i,j,l} - \omega_{SE,j,l}\right) +}_{Occupational\ composition} \end{aligned} \tag{9}$$

Hence, the above decomposition shows that the difference between two GEIs can be decomposed between an effect coming from the difference in sector indices, a second effect from the differences in the sectoral composition (e.g. higher share of employment in industry) and a third effect capturing the differences across MS in terms of occupational composition. There is no occupation index effect due to the fact that these indices do not vary across MS.

The three quantities defined above can be computed for each time period t and are then averaged. These average values are displayed in Figure 10. BG and SI were not included in this first version of this exercise as it would have required converting the 3-digit occupations for SE in terms of 2-digit codes for these two MS.

Figure 10: GEI decomposition



Source: Own elaboration.
Note: Decomposition of the point difference between the GEI of each MS and SE's GEI. The decomposition includes three components. One coming from differences in the sector index ('sector index'), a second contribution from the difference in the sectoral composition of employment ('sector'), and a last contribution from occupational employment composition ('occup'). 'diff GEI' shows the point difference between the GEI and SE's GEI. Numbers are averaged over the period 2011-2022 and ordered according to the 'sector index' contribution.

The results in Figure 10 indicate that the bulk of the difference in GEIs originates from the sector index. This highlighted by the fact that the ranking of MS in terms of the sector index contribution (from smallest to largest in Figure 10) almost perfectly reproduce the GEI ranking. A few points are nevertheless worth noting.

First, the occupational composition of employment contributes positively[32] to the GEI of LV (+1.2 point on average), a point also valid for MS found at a lower end of the GEI ranking (e.g. RO, PL). A natural explanation would be related to the agricultural (as already discussed previously), though this deserved to be confirmed for LV. Second, the occupational composition can also have negative

[32] When compared to the occupational composition of employment in SE. The use of SE as a reference is implicit when discussing the results from this exercise.

contributions in few MS (-1.2 points in DK and CY, -1 point in IT) meaning that the GEI would have been slightly greater on average in these MS if their occupational composition matched SE's one. The sector contribution is negligeable for most MS with the exception of PT, RO and PL (respectively, -1.6 points, -1.2 points and -0.9 point).

Therefore, these results indicate that the sector pillar is central to understanding MS variations in their GEI. The sector index and its (sub)dimensions appear as key drivers for the greening of employment, much more than changes in the sectoral and occupational composition of employment. These results can be seen as indicative of negligeable effects from sectoral/occupational shift in employment on the greening of employment, though they cannot be considered definitive given the high level at which data (particularly sectors) is aggregated (i.e. some employment reallocations are likely taking place within 1-digit NACE groups).

Finally, it is worth mentioning that these results should be extended to better understand the sector index contributions and the role of the different (sub)domains. Some of the results discussed above would probably deserve additional investigations (e.g. for LV, PT).

CONCLUSION

This study presents a novel attempt to develop a composite indicator to measure the greenness of jobs across EU Member States. The Green Job Index (GJI) and the Green Employment Index (GEI) provide an integrated framework that incorporates sectoral and occupational approaches to measuring green employment. By aggregating multiple dimensions—including energy use, emissions, waste generation, task content, and skills—this study offers a different perspective on the greening of EU labour markets currently taking place.

The methodology follows best practices in composite indicator construction, drawing upon the framework proposed by Nardo et al. (2005). The study adopts a hierarchical structure that organises green employment characteristics into two main pillars – sectoral (production-based) and occupational (job content-based) – each further divided into input, process, and output dimensions. To integrate the diverse set of indicators, a combination of Principal Component Analysis (PCA) and Dynamic Factor Models (DFM) is used to extract common trends and construct indices. This approach ensures that the composite indicator is data-driven, while trying to maintain transparency and interpretability. The resulting indices offer a novel way to quantify employment greenness in a multidimensional manner, avoiding binary classifications that, although very relevant from a policy standpoint, may oversimplify the complexity of the green job concept.

Analysis of the GEI for 2022 reveals that SE recorded the highest level of employment greenness, followed by FI, LV and AT. In contrast, HR, EE and EL exhibited the lowest GEI values. The evolution of the GEI between 2011 and 2022 further shows that employment greenness has increased in nearly all but one Member States, with DK experiencing the most notable improvement. Interestingly, countries with initially lower GEI values have, on average, recorded higher increases, indicating a trend toward convergence in employment greenness across the EU.

In addition, a little exercise was performed at the end of the paper to showcase some of the possible analysis using the GEI. Cross Member States variations in the GEI were decomposed into the contribution originating from variations in the sector indices, and in the sectoral and occupational composition of employment. The results revealed that the different sectoral and occupational distributions between MS only accounted for a fraction of the differences in the GEI. This result highlighting the key role of improvements in input, process and output at sectoral level to drive the greening of employment in the EU.

Overall, this study tries to contribute to the ongoing efforts around the operationalisation of the green job concept. By providing a continuous and multidimensional assessment of job greenness, the GJI and GEI offer complementary information on the ongoing progress towards greener employment. This evidence could prove useful for policymakers aiming to monitor and promote the transition to a greener economy, though future refinements, primarily focused on methodological considerations, are required to enhanced the robustness of the GEI.

References

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings*. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (Vol. 4, pp. 1043–1171). Elsevier.

ANPAL. (2023). *La vocazione Green delle professioni Un indice di Green Rating a partire dalla classificazione ESCO* (Working Paper Series). ANPAL.

Apostel, A., & Barslund, M. (2024). Measuring and characterising green jobs: A literature review. *Energy Research & Social Science*, *111*, 103477.

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, *103*(5), 1553–1597.

Bachelot, M. (2024a). *Using O* NET green jobs and tasks in Europe? A critical assessment based on French data*. https://cnam.hal.science/hal-04652272/

Bachelot, M. (2024b). *Using O* NET green jobs and tasks in Europe? A critical assessment based on French data*.

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

Bańbura, M., & Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with aribitrary pattern of missing data. *Journal of Applied Econometrics*, *29*(1), 133–160.

Bluedorn, J., Hansen, N.-J., Noureldin, D., Shibata, I., & Tavares, M. M. (2023). Transitioning to a greener labor market: Cross-country evidence from microdata. *Energy Economics*, *126*, 106836.

Bohnenberger, K. (2022). Is it a green or brown job? A Taxonomy of Sustainable Employment. *Ecological Economics*, *200*, 107469.

Bontadini, F., & Vona, F. (2023). Anatomy of Green Specialisation: Evidence from EU Production Data, 1995–2015. *Environmental and Resource Economics*, *85*(3–4), 707–740. https://doi.org/10.1007/s10640-023-00781-7

Bowen, A., Duffy, C., & Fankhauser, S. (2016). Green growth'and the new industrial revolution. *Policy Brief. Grantham Research Institute on Climate Change and the Environment and Global Green Growth Institute*. https://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2016/01/Bowen-et-al-2016.pdf

Bowen, A., & Hancké, B. (2019). The Social Dimensions of" Greening the Economy. *European Commission*. https://www.bollettinoadapt.it/wp-content/uploads/2020/01/KE0219988ENN.en_.pdf

Bowen, A., & Kuralbayeva, K. (2015). Looking for green jobs: The impact of green growth on employment. *Grantham Research Institute Working Policy Report. London: London School of Economics and Political Science*, 1–28.

Caprile, M., Sanz, P., Riobóo, I., Welz, C., & Rodriguez, R. (2017). *Mapping Varieties of Industrial Relations: Eurofound's Analytical Framework Applied*.

Cascales Mira, M. (2021). New Model for Measuring Job Quality: Developing an European Intrinsic Job Quality Index (EIJQI). *Social Indicators Research*, *155*(2), 625–645.

Castillo, M. (2023). *Green jobs, green economy, just transition and related concepts: A review of definitions developed through intergovernmental processes and international organizations*. International Labour Office.

Chen, C., & Liu, L.-M. (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, *88*(421), 284–297. https://doi.org/10.1080/01621459.1993.10594321

Consoli, D., Marin, G., Marzucchi, A., & Vona, F. (2015). Do Green Jobs Differ from Non-Green Jobs in Terms of Skills and Human Capital? *SSRN Electronic Journal*.

Dierdorff, E. C., Norton, J. J., Drewes, D. W., Rivkin, D., & Lewis, P. (2009). *Greening of the World of Work: Implications for O*NET®-SOC and New and Emerging Occupations*.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi–maximum likelihood approach for large, approximate dynamic factor models. *Review of Economics and Statistics*, *94*(4), 1014–1024.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (Vol. 38). OUP Oxford.

Dutta, S., Lanvin, B., Wunsch-Vincent, S., & León, L. R. (2022). *Global innovation index 2022: What is the future of innovation-driven growth?* (Vol. 2000). WIPO.

EC: Directorate-General for Research and Innovation. (2024). *EU eco-innovation index 2024*. Publications Office of the European Union.

ESCO. (2022). *Green Skills and Knowledge Concepts: Labelling the ESCO classification*. European Commission.

European Institute for Gender Equality. (2017). *Gender equality index 2017: Methodological report*. Publications Office.

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *The Review of Economics and Statistics*, *89*(1), 118–133.

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Social Indicators Research*, *141*(1), 61–94.

Greenacre, M., Groenen, P. J., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1), 100.

Hsu, A., & Zomer, A. (2016). Environmental Performance Index. *Wiley StatsRef: Statistics Reference Online*, 1–5.

Juwana, I., Muttil, N., & Perera, B. J. C. (2012). Indicator-based water sustainability assessment—A review. *Science of The Total Environment*, *438*, 357–371.

Kim, D., Kim, Y., Voos, P., Suzuki, H., & Kim, Y. D. (2015). Evaluating Industrial Relations Systems of OECD Countries from 1993 to 2005: A Two-Dimensional Approach. *British Journal of Industrial Relations*, *53*(4), 645–663.

Leschke, J., & Watt, A. (2014). Challenges in Constructing a Multi-dimensional European Job Quality Index. *Social Indicators Research*, *118*(1), 1–31.

Madurai Elavarasan, R., Pugazhendhi, R., Irfan, M., Mihet-Popa, L., Campana, P. E., & Khan, I. A. (2022). A novel Sustainable Development Goal 7 composite index as the paradigm for energy sustainability assessment: A case study from Europe. *Applied Energy*, *307*, 118173.

Maldonado, J. E., Vandeplas, A., Vanyolos, I., Vigani, M., & Turrini, A. (2024). *Assessing Green Job Dynamics in the EU A Comparison of Alternative Methodologies*. Directorate General Economic and Financial Affairs (DG ECFIN), European ….

Mariano, R. S., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *18*(4), 427–443.

Mori, K., & Christodoulou, A. (2012). Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review*, *32*(1), 94–106.

Munda, G. (2012). Choosing Aggregation Rules for Composite Indicators. *Social Indicators Research*, *109*(3), 337–354.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). Handbook on constructing composite indicators: Methodology and user guide. *OECD Statistics Working Papers 2005/03*.

Nussbaumer, P., Fuso Nerini, F., Onyeji, I., & Howells, M. (2013). Global insights based on the multidimensional energy poverty index (MEPI). *Sustainability*, *5*(5), 2060–2076.

OECD. (2024). *OECD Employment Outlook 2024: The Net-Zero Transition and the Labour Market*. OECD Publishing,.

Papadimitriou, E., Norlen, H., & Del, S. M. (2020). *JRC Statistical Audit of the 2020 Gender Equality Index*. JRC Publications Repository. https://doi.org/10.2760/39645

Poncela, P., Ruiz, E., & Miranda, K. (2021). Factor extraction using Kalman filter and smoothing: This is not just another survey. *International Journal of Forecasting*, *37*(4), 1399–1425.

Shumway, R. H., & Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, *3*(4), 253–264.

Sommers, D. (2013). BLS green jobs overview. *Monthly Lab. Rev.*, *136*, 3.

Stanef-Puică, M.-R., Badea, L., Şerban-Oprescu, G.-L., Şerban-Oprescu, A.-T., Frâncu, L.-G., & Crețu, A. (2022). Green jobs—A literature review. *International Journal of Environmental Research and Public Health*, *19*(13), 7998.

Stock, J. H. (1989). New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual/MIT Press*.

Stock, J. H., & Watson, M. W. (2008). *Business Cycles, Indicators, and Forecasting*. University of Chicago Press.

Stoevska, V., & Hunter, D. (2012). Proposals for the statistical definition and measurement of green jobs. *International Labour Office Geneva*. http://old.adapt.it/adapt-indice-a-z/wp-content/uploads/2014/05/stoevska_hunter.pdf

UNEP/ILO/IOE/ITUC. (2008). *Green Jobs: Towards decent work in a sustainable, low-carbon world*.

Urban, P., Rizos, V., Ounnas, A., Kassab, A., & Kalantaryan, H. (2023). *Jobs for the green transition: Definitions, classifications and emerging trends*.

Valero, A., Li, J., Muller, S., Riom, C., Nguyen-Tien, V., & Draca, M. (2021). *Are'green'jobs good jobs?* Centre for Economic Performance, LSE.

Vona, F. (2021). *Labour markets and the green transition: A practitioner's guide to the task based approach* (Vol. 126681). Publications Office of the European Union. https://air.unimi.it/handle/2434/905262

Vona, F., Marin, G., Consoli, D., & Popp, D. (2018). Environmental Regulation and Green Skills: An Empirical Exploration. *Journal of the Association of Environmental and Resource Economists*, *5*(4), 713–753.

World Health Organization. (2024). *The World Health Organization-Five Well-Being Index (WHO-5)*. orld Health Organization.

ANNEX

A. NARDO ET AL., (2005)'S METHODOLOGY

The 10 steps are reproduced from Nardo et al., (2005) and are the following:

1. Developing a theoretical framework that defines and structures what is measured and provides the basis for the selection and combination of variables into a meaningful index.
2. Selecting variables based on the analytical soundness, measurability, country coverage, cross-country comparability, and relevance of indicators.
3. Imputing missing data in order to obtain a complete dataset for all countries.
4. Conducting a multivariate analysis to study the overall structure of the dataset, assess its suitability, and guide subsequent methodological choices.
5. Normalising the data, if needed, to ensure the comparability of variables.
6. Weighting and aggregating indicators according to both the theoretical framework and the results of the multivariate analysis.
7. Conducting an uncertainty and sensitivity analysis to assess the robustness of the index in terms of all possible sources of uncertainty in its development (choice of imputation method, normalisation scheme, weighting system or aggregation method).
8. Returning to the data in order to analyse what domains and sub-domains are driving the index results.
9. Identifying possible association with other variables and existing known indicators.
10. Presenting and disseminating the index results in a clear and accurate manner.

B. Data

B1. Indicators

B1.1. Sector – Inputs

<u>*Energy*</u>

Data on energy is retrieved from the Physical Energy Flow Accounts (PEFA) published by Eurostat. PEFA provides information on energy flows across economic sectors (and households). These flows are split between energy flows from the environment to the economy; within the economy (i.e. energy used as intermediate input or for final consumption); and back to the environment from the economy (usually in the form of residual heat). PEFA further distinguishes between the 'supply' and 'use' of energy, classified into 36 categories corresponding broadly to energy sources. The 'use' of energy is relevant to obtain a measure of energy entering the economy as an input. Among the 36 categories, two broad groups are considered for the energy subdimension. These are natural resources (e.g. coal, wind) used as inputs to produce energy products (e.g. electricity) and energy products used as inputs within the economy. Natural resources are classified into renewable (e.g. hydro based, solar based) and non renewable (e.g. from fossil origin), and two energy products, namely liquid biofuels and biogas, can be considered as green energy products, whereas all remaining products are considered as non-green products[33]. One important exception (and limitation of PEFA) regards electricity for which it is not possible to disentangle electricity originating from renewable and non-renewable natural resources. For this reason, electricity is not used to create indicators for the energy dimension.

The series from PEFA are used to create five indicators. The first one is a measure of energy productivity (i.e. the inverse of energy intensity) computed as the sum of all natural resources and energy products (including electricity) measured in terajoule at the sector level, divided by sectoral value added (VA)[34]. The other four indicators correspond to the shares of different energy sources

[33] Nongreen energy products are the products ranging from hard coal [P08] to nuclear fuel [P22]. Furthermore, the product [P23] wood, wood waste and other solid biomass, charcoal could be considered a green energy product but was not in this version of the GJI.

[34] Obtained from Eurostat [nama_10_a64].

expressed in terms of the total use of energy[35]. Two indicators are related to renewable and non-renewable natural resources. Natural resources are used as inputs to create energy products and are only relevant for few sectors (primarily sector D - Electricity, gas, steam and air conditioning supply). As a result, these two indicators are created at MS level and not at the sectoral level. These indicators should nonetheless inform (imperfectly) on the overall greenness of energy products used in the entire economy. The last two indicators are computed as the share of green and non-green energy products used as inputs at the sectoral level. Indicators related to non-renewable and non-green energy products are expressed as 1 minus the indicator (e.g. 1 minus the share of non-green products).

Table 1 presents a summary of indicators used for the energy subdimension and Figure 11 to Figure 18 in Annex B2.1 display the standardised indicators (i.e. with zero mean and unit variance) by MS and by NACE sectors[36].

Without delving too much into the evolution of each indicator, these figures reveal an overall trend over the sample period consistent with the idea that energy inputs have been greening across MS and sectors. This is particularly true at MS level for indicators related to the shares of green, renewable and non-renewable energy inputs (Figure 13, Figure 17 and Figure 18) and to a lower extent for energy productivity and the share of non-green products, though a positive trend is still observed in a majority of MS (Figure 11 and Figure 15).

At sector level, the share of green products has increased between 2014 and 2021 in all sectors with the exception of Agriculture, fishing and forestry (Figure 14). For energy productivity, and the share of non-green products, the evolution across sectors is more heterogenous (Figure 12 and Figure 16).

Capital

Three different data sources are used to create indicators for the capital subdimension.

The first one is the Environmental Protection Expenditure Accounts (EPEA). EPEA tracks economic resources allocated to activities and products designed to protect the environment and mitigate climate change. In particular, EPEA provides information on investment in climate change

[35] Taking the sum of natural resources and energy products might not be correct from a conceptual perspective as these two categories correspond to two different flows. Nevertheless, this provide a normalising constant to express and compare these indicators across MS.

[36] Standardised indicators are displayed as these series will be used to construct the various indices.

mitigation, which should be understood as the allocation of financial resources to activities designed to reduce or prevent GHG emissions. Data for sector B is not collected and the sectors G to U are aggregated into one broad service sector, except for sectors H, J and M, for which data is available separately. However, when trying to compute investment for the service sector net of sectors H, J and M, the series are often negative. For this reason, it is assumed that investment in service sectors outside H, J and M is equal to zero. Moreover, information for sector A is missing in many MS, though the available data reveals that investment in this sector tend to be small (in terms of value added). Thus, investment in sector A is assumed to be zero for all MS[37]. Investment data for sector E is also always equal to zero. An indicator is created by dividing investment with sectoral VA retrieved from Eurostat.

The second indicator is constructed using information on green bonds. Green bonds are used to finance activities that address climate change and environmental issues and should enable green investment. The data is obtained from the European Environment Agency and is calculated based on the issuance of green bonds by companies and governments in the EU excluding supranational entities. The series is provided as a share of total bonds issuance and is used directly as an indicator without transformations. The data starts in 2014 and shows that BG, CY and CZ have not issued any green bonds over the sample period. Some care should be taken with this data as it is collected from different data providers in different MS[38]. However, this indicator is included within the monitoring framework for the 8th Environment Action Programme[39], which supports its inclusion within the framework for the computation of the composite indicator.

The final indicator comes from the Eco-Innovation Index, developed by the European Union to measures the performance of EU countries in terms of eco-innovation and captures the progress in the transition to a green and sustainable economy. It evaluates countries based on their ability to develop and adopt innovative solutions that reduce environmental impacts, enhance resource efficiency, and support sustainable growth. The index is composed of five dimensions: eco-

[37] This is a simplifying assumption that could be relaxed in future versions of the index.

[38] See https://www.eea.europa.eu/en/analysis/indicators/green-bonds-8th-eap.

[39] See https://eur-lex.europa.eu/legal-content/?uri=COM%3A2022%3A357%3AFIN.

innovation inputs, eco-innovation activities, eco-innovation outputs, resource efficiency outcomes, and socio-economic outcomes.

The eco-innovation input dimension is used as an indicator for the capital subdimension. Eco-innovation inputs are related to R&D expenditures in green technologies, economic support for eco-innovation in the form of financial mechanisms (e.g. green funds) and access to skilled human capital (e.g. environmental scientists, engineers). The later dimension is more relevant regarding the labour input but intuitively, a higher index should be associated nonetheless with a greener capital in the MS overall economy. Data is available for all MS starting from 2013.

These three indicators inform (indirectly) on the greenness of the capital inputs to the production process given that a higher investment in climate mitigation and/or share of green bonds and/or performance in terms of eco-innovation input should be associated with a greener capital. It should be noted that the eco-innovation index and the share of green bonds are only available at MS level, whilst information on investment is only available for a subset of sectors (i.e. C, D, F, H, J and M). Therefore, the capital subdimension will only provide limited sectoral information and will be more relevant to compare MS.

Figure 19 in the annex shows that investment in climate mitigation only increased in a handful of MS over the 2006-2021 time period (i.e. DE, FI, FR, HU, LT and SE). At sectoral level (Figure 20), investment as a share of VA has increased in sector C, and has been stable or decreased in the remaining sectors, in particular, sectors F and H (respectively construction and transport). With regards to green bonds, their issuance has increased in a majority of MS, particularly in most recent times (Figure 21). Likewise, the indicator based on the eco-innovation input displays improvements in most MS (Figure 22).

Land use

Land use data can be obtained from the Land Use and Coverage Area Frame Survey (LUCAS), which provides harmonised and detailed information on land use and land cover. The survey is conducted every three years, and it categorises areas based on their purpose, such as agriculture, forestry, or residential. This data is provided by economic activity according to its own classification. LUCAS sectors can be matched to 1-digit NACE sectors, except for some activities that require a slight adjustment discussed. Data is available between 2009 and 2018[40]. Given that the survey is run every three years, only four observations are available per MS and sectors in total. As a result, this indicator should be treated with care, though its inclusion is relevant from the perspective of the preservation and restoration of biodiversity. The land use indicator is created by computing the share of land use by activities expressed in terms of the total land area[41] and is expressed as 1 minus the share of land use to ensure that an increase implies that the activity covers less land space.

Figure 23 and Figure 24 display the indicator at MS and sectoral level. The indicator in most MS shows a hump-shaped evolution between 2009 and 2018, usually increasing until 2015 (i.e. less land used across economic activities on average) and decreasing in 2018. The indicator constantly improved over the period in BG, IT and LV.

At sector level, Figure 24 shows that land use for some activities usually considered to have heavy environmental impacts (i.e. A, B and C) has decreased on average across MS. On the other hand, land use for service activities appear to have increased, in particular in sector H, I, J, M and N.

Though not visible from both figures, it is worth mentioning that the vast majority of land is used for agricultural purposes (around 70% of land use on average), with the remaining activities accounting for a marginal share of the total land area (usually less than one percent).

B1.2. Occupation - Input

<u>Labour</u>

Indicators on skills are created from the European Skills, Competences, Qualifications, and Occupations (ESCO) database. ESCO provides information on more than 13,000 skills and knowledges assigned across 3,000 occupations. In ESCO, a knowledge is defined as a 'body of facts, principles, theories and practices' (ESCO, 2018) and a skill consists in the 'ability to apply knowledge and use know-how to complete tasks and solve problems'. The ESCO occupation classification expands on the 4-digit ISCO classification and is therefore much finer. Hence, it will be required to aggregate the data back to the 3-digit ISCO level given that skills and knowledges are assigned to ESCO occupations only above the 4-digit level.

[40] A new release for 2022 should soon become available.

[41] Available from Eurostat [reg_area3].

ESCO provides its own classification of green skills defined as ‘the knowledge, abilities, values, and attitudes required to support environmental sustainability across occupations and industries’ (ESCO, 2022). In total, there are 570 skills and knowledges labelled as green[42]. Skill and knowledge are further classified into essential and non-essential, where non-essential should be understood as skill relevant only for a subset of jobs within the occupation (ESCO, 2018).

From the ESCO database, I construct an indicator which consists in the share of green skills and knowledges (or concepts) among the total number of concepts per occupation. Two additional indicators are constructed from the share of essential skills and essential knowledges expressed in terms of the total number of skills/knowledges. The final indicators used to compute the skill subdimension are further transformed as $log(1 + \text{the indicator})$ since the share of green skills/knowledges is large for a couple of 3-digit occupations (e.g. 621 – Forestry and Related Workers and 961 – refuse workers) and this transformation allows for ‘smoothing’ larger values (see Figure 47).

Before computing these indicators, the ESCO dataset needs to be cleaned as skills can appear as duplicates[43]. In ESCO, skills are classified into eight major groups[44] (e.g. management skills, constructing skills) and the same skill, identified by its uniform resource identifier (uri), can often appear in different categories. This issue should not be really problematic when one only cares about the skills since the duplicates can simply be dropped. However, this is not necessarily the case when the mapping to occupations matters and simply dropping duplicates could lead to a loss of information.

Figure 46 and Figure 47 display the green skill/knowledge indicators in their transformed and original form, respectively. Figure 48 and Figure 49 complement this evidence with numbers on the share of green jobs obtained according to these different indicators in the EU and at MS level.

---

[42] For unknown reasons, this number is lower when retrieving this information from the ESCO API with 512 green knowledges/skills available. The large majority (480) is present in both data but there are 32 (90) skills only available from the ESCO API (ESCO csv download). The indicators are computed based on 602 unique skills available by merging both datasets.

[43] Duplicate skills appear when retrieving data from the ESCO API. It is unclear how these duplicates are addressed in the data directly provided by ESCO as csv files.

[44] Knowledges are classified into 12 groups.

These numbers can be compared with available evidence obtained using ESCO (Maldonado et al. 2024) to analyse the potential effects from the different methodologies used to construct indicators. Maldonado et al. (2024) use a slightly different indicator based on skills only (both essential and optional). Nevertheless, the results across the studies are mostly consistent as the indicator based on both concepts leads to 11 3-digit ISCO occupations with more than 10% green concepts, while Maldonado et al. (2024) report 14 occupations with more than 10% green skills, 10 of which overlap[45]. In terms of employment at EU level, both studies converge around a share of green jobs close to 7% in 2022, with this share decreasing over the last decade[46]. Similar results are also obtained at MS level (results not reported) with a few exceptions (FR, IE, IT ranked lower in this study and BG, HU and SI higher).

Regarding the other two indicators based on essential knowledge and skills, it is worth noting that green essential knowledge represents a much higher of share essential knowledge when compared to essential green skills (Figure 47)[47]. This is reflected in the different share of green jobs resulting from these indicators with essential knowledges leading a much greater share (between 27% and 29% in Figure 48). The share of green jobs obtained from essential green skills lies slightly below 5%, a number similar to task approach when efforts are made to account for aggregation problems (Vona, 2021).

B1.3. Sector - Process

Waste and circularity

Waste Generation Statistics provide comprehensive data on the types and quantities of waste produced across the European Union. This dataset is used for monitoring waste management policies, assessing environmental pressures, and driving circular economy initiatives. Therefore, indicators

---

[45] Occupations 631 and 633 are not included in the ESCO sample retrieved from the API, which calls for verifications on why this is the case. Maldonado et al. (2024) find that these two occupations are green based on their indicator.

[46] Maldonado et al. (2024) show that the main factor driving the decrease comes from the lower share of employment in the agricultural sector (e.g. in EL, PL and RO).

[47] On average green essential knowledge represent 8.3% of essential knowledges with a median of 3.6%. The same figures for essential green skills are respectively 3.1% and 1.2%.

based on this dataset should be closely related to the environmental objectives that are part of the green job definition.

The dataset covers waste generation by different actors (e.g. industries, households), offering a detailed breakdown by waste categories and sectors. It is collected under the EU Waste Statistics Regulation, ensuring harmonisation and comparability across MS. This regulation categorises waste by for more than 30 material types (e.g. metals, plastics) available on Eurostat's website. These 30 types are further categorised in broader groups such as recyclable waste, chemical and mineral wastes. A special attention is also given to hazardous waste due to its potential environmental and health risks.

Data on waste generation is used to construct four indicators. These include series on VA per one tonne of waste for each available sectors, and the share of recyclable, chemical and hazardous wastes among total sectoral wastes. The last two indicators are expressed as $1 -$ the relevant shares.

Waste generation statistics are collected on bi-annual basis and data is available between 2006 and 2022. The service sector is aggregated for sectors G to U.

An additional indicator for this subdimension is obtained from the circularity material use rate (CMUR). The CMUR represents the proportion of material resources used in an economy that come from recycled or reused materials. It is computed as the amount of secondary material used (i.e. Materials recovered and reintroduced into the economy through recycling or reuse) over the total amount of material used.

This indicator is only available at MS level for the period 2010-2023.

Figure 25 to Figure 33 display these indicators by MS and sectors. Waste generation as improved in some MS, (i.e. AT, BG, ES, FR, HU, SE and RO) with other showing no evolution or a decrease in the indicator. At sector level (Figure 26), the average indicators for sector C (manufacturing) and services (G-S) have increased over the time period. By type of waste, the evolution appear more contrasted with substantial differences across type of waste, MS and sectors. The indicator related to chemical waste is displayed in Figure 27 and Figure 28, in Figure 29 and Figure 30 for hazardous waste, and Figure 31 and Figure 32 for recyclable waste. The CMUR has increased in a majority of MS between 2010 and 2023 (Figure 33).

Air emissions

Air emission data originates from Air Emission Accounts (AEA), which measure air emissions in physical terms (e.g., tonnes) and assigns them to economic activities. Unlike territorial emission inventories, AEA takes a production-based approach, assigning emissions to economic sectors based on their activities, regardless of where emissions occur geographically (resident approach). The dataset provides annual data and is harmonised across MS, ensuring consistency and comparability over time and across countries.

AEA includes a wide range of air pollutants, such as greenhouse gases (e.g., CO_2, CH_4, N_2O), particulate matter (PM10, PM2.5) and nitrogen oxides (NO_x). These air pollutants are aggregated by Eurostat into three categories based on their different negative effects on the environment. These categories are 'Global warming potential' which includes GHG, 'Acidifying gases' (ACG) and 'Tropospheric ozone precursors' (O3PR).

These three categories together with the series on particulate matter smaller than 2.5 micron are used to create four indicators expressed as sectoral VA per one tonne of emission (the inverse of emission intensities).

Data on air emissions is available between 2006 (for few MS) and 2022.

The figures in Annex B2.2 indicate that air emissions of the various pollutants tend to have improved in most MS when expressed in terms of VA. This is particularly true for GHG emissions (Figure 36) but also concerns the three other indicator (Figure 34, Figure 38, and Figure 40 for ACG, O3PR and PM2.5). At sectoral level, the indicators increase for all sectors with the exception of sector B (mining and quarrying).

B1.4. Occupation - Process

Tasks

Information on tasks is obtained from the Occupational Information Network (O*NET) database which provide detailed information on the attributes of occupations across the U.S. economy. O*NET is the reference source of information on tasks and as such it is a well-known database (see Bachelot, 2024 for instance). O*NET collects data on tasks (i.e. job specific actions and duties) divided into core and supplemental tasks. In general, each occupation is assigned between 10 and 25 tasks.

O*NET provides information on green tasks defined as '*actions, processes, or responsibilities that contribute to reducing environmental impact, promoting energy efficiency and the use of renewable energy, enhancing waste reduction, recycling, and resource management, and supporting climate adaptation and mitigation efforts'*. Green tasks are not limited to traditionally green industries but can be found across all sectors.

The occupational classification used in O*NET builds on the Standard Occupation Classification (SOC) of the Bureau of Labor Statistics and does not corresponds perfectly to the ISCO classification used in Europe. It is therefore necessary to use a crosswalk to convert O*NET-SOC codes into ISCO ones. This step is important and is likely to lead to an important loss of precision when computing the indicator. Furthermore, it appears that O*NET is no longer updating its list of green tasks[48], which relies on the previous (i.e. 2010) SOC classification. As a result, I use the crosswalk provided by O*NET to convert SOC-2010 codes into the SOC-2019 ones. At this stage, I further perform a simple word search and identify around 150 additional tasks that could be considered green. To try addressing concerns related to the future aggregation at 3-digit ISCO codes, 8-digit O*NET-SOC codes are dropped from the sample and only 6-digit codes are converted to ISCO codes[49]. These 8-digit O*NET-SOC measure occupations at a very detailed level and dropping these will avoid aggregating occupations at a level, which is much finer than the 3-digit ISCO occupation.

The conversion to ISCO codes makes use of the crosswalk developed by ESCO. This crosswalk presents the advantage of informing on the precision of the match between an ISCO and a O*NET-SOC codes. The categories include exact (ISCO) match, close match, broad match and narrow match. Only exact and close matches are considered when converting O*NET-SOC codes to ISCO codes.

Following these adjustments, an indicator of green tasks is obtained by computing the share of green tasks among the total number of tasks by occupation.

B1.5. Sector – Outputs

Indicators for the green production subdimension are obtained from three different data sources. First, from the well-known environmental goods and services sector (EGSS) dataset, provided by Eurostat.

---

[48] When retrieving data from the API, information on green tasks is no longer provided.

[49] Actually, most of the 8-digit O*NET-SOC codes were dropped as part of the change from to the 2010 to 2019 SOC classification.

This dataset is meant to measure the contribution of the green economy to the overall economy and considers as green, economic activities dedicated to environmental protection aimed at preventing or mitigating pollution and conserving ecosystems, and resource management activities concerned with with the sustainable use and management of natural resources. This includes industries involved in renewable energy production, energy efficiency technologies, waste management, recycling, sustainable agriculture, and environmental consulting, among others.

These two categories of activities could be used to create two indicators, but for this version of the GJI, only one indicator was created based on the total green output by sectors and expressed as a share of VA. The data is available as early as 2006 (only for the NL) and series for most MS become available around 2013-2014. The series currently run until 2022.

In addition to the EGSS, I use two sources already used to generate indicators for the input dimension. More precisely, the eco-innovation index (EC: Directorate-General for Research and Innovation, 2024) also includes an eco-innovation output dimension. This dimension is meant to capture the results of eco-innovation activities, and reflect the extent to which green innovations translate into practical applications, technologies, and market success. Indicators for this dimension include eco-innovative patents, the availability and uptake of eco-friendly products and services, the publications and knowledge sharing linked to eco-innovations and the export of eco-innovative goods and services. As for the eco-innovation input (Section B1.1), this index is only available at MS level.

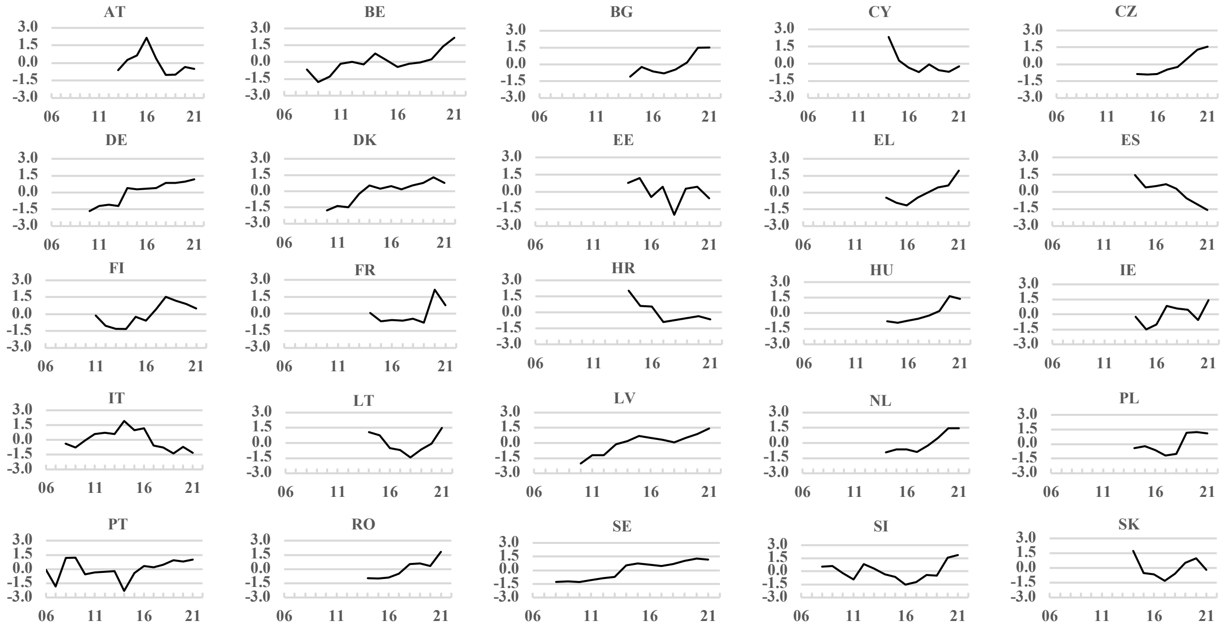
A last indicator is constructed relying the Physical Energy Flow Account used for the energy subdimension. Instead of relying on the 'use' category, I make use of the 'supply' category to compute the share of non-green energy products among energy products supplied. Non-green energy products are those identified to generate the indicator on non-green energy product for the energy subdimension.

B2. Figures

B2.1. Sector – Inputs

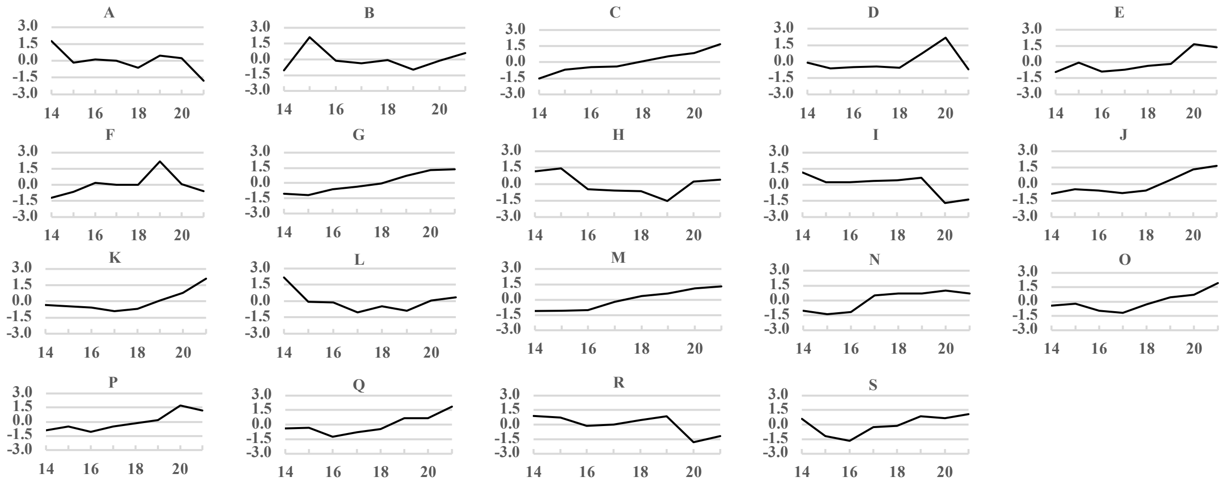
Energy

Figure 11: Energy productivity – Member States



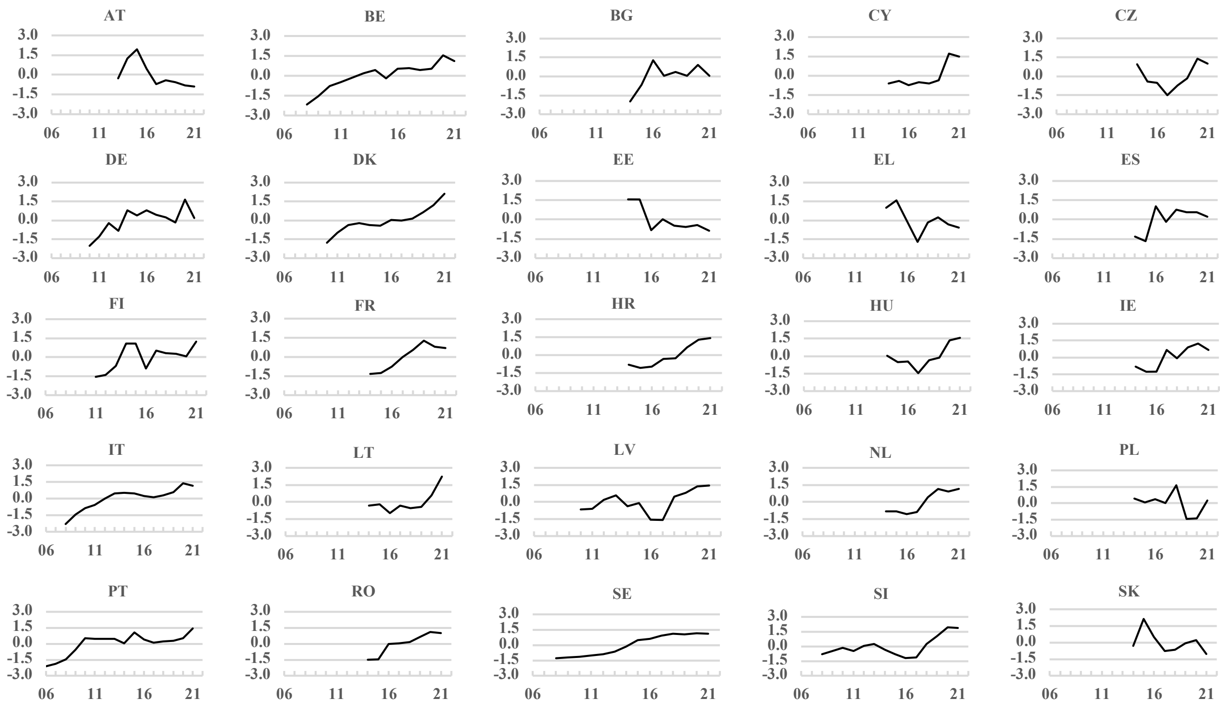
Source: Eurostat [env_ac_pefasu].
Note: Energy productivity is computed as the ratio of sectoral value added over the sum of energy used. Series are standardised to have zero mean and unit variance. The series at MS level are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1 for additional details.

Figure 12: Energy productivity – Sector



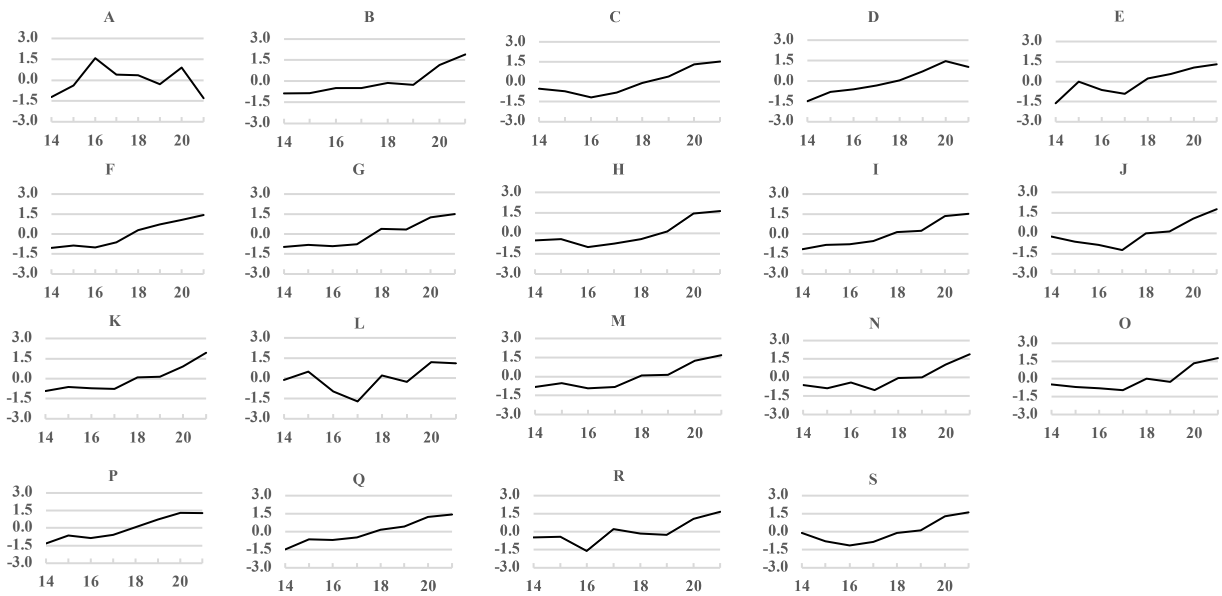
Source: Eurostat [env_ac_pefasu].
Note: Energy productivity is computed as the ratio of sectoral value added over the sum of energy used. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. 2014 is the first year for which all MS provide data. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 13: Share of green energy products – Member States



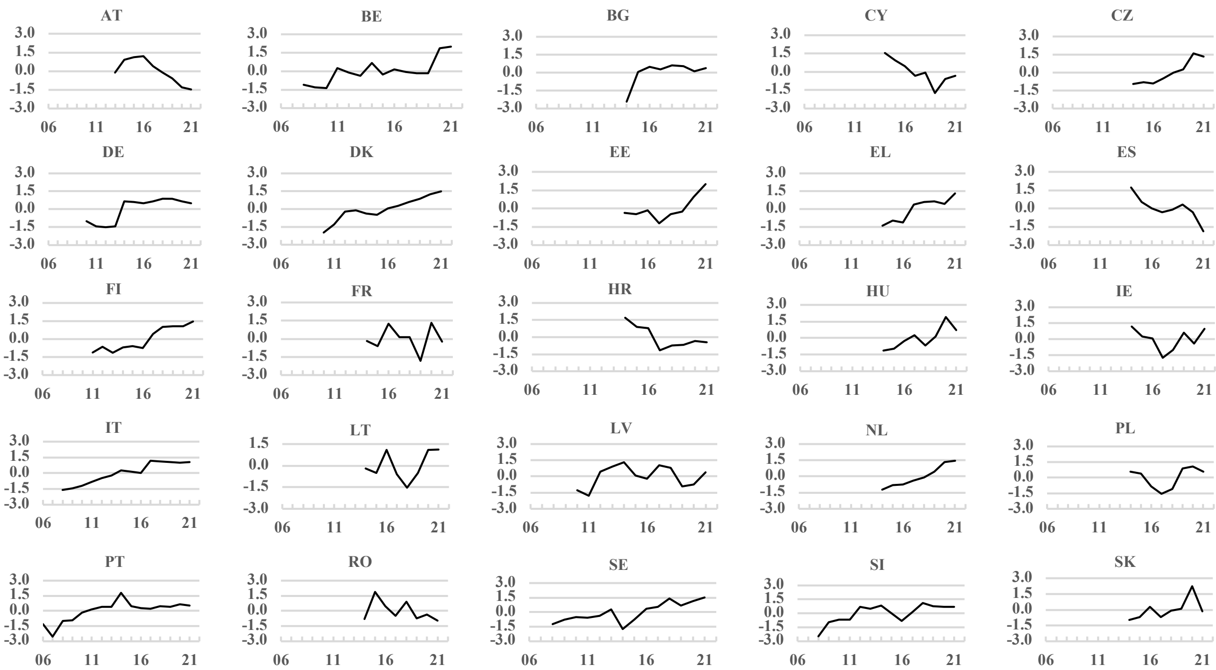
Source: Eurostat [env_ac_pefasu].
Note: Green products correspond to liquid biofuels [P24] and biogas [P25]. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1 for additional details.

Figure 14: Share of green energy products – Sectors



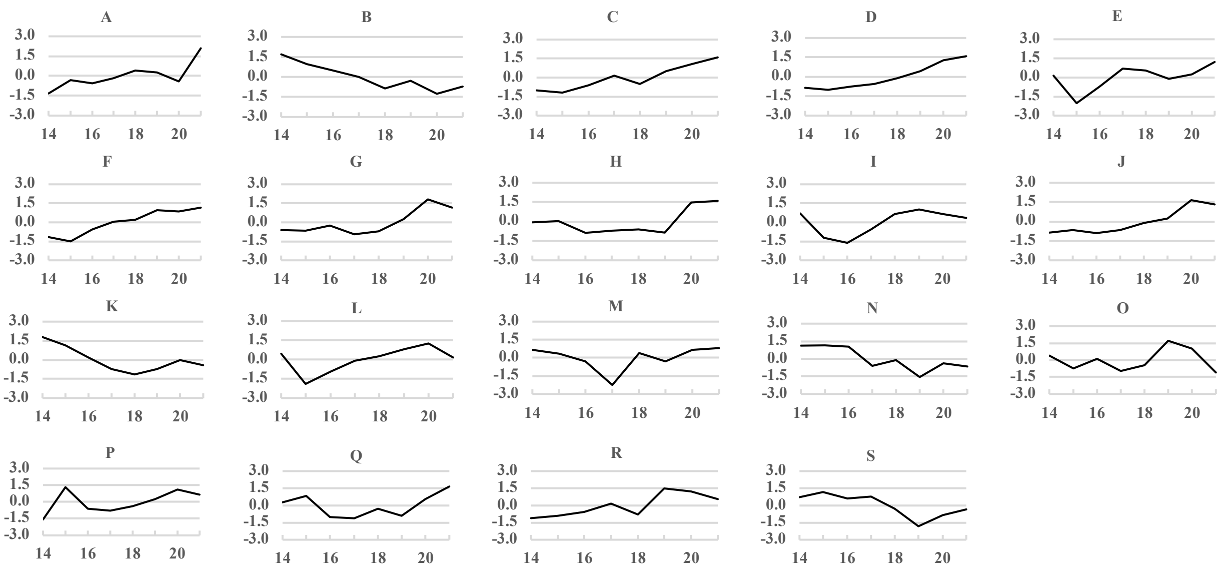
Source: Eurostat [env_ac_pefasu].
Note: Green products correspond to liquid biofuels [P24] and biogas [P25]. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 15: Share of non-green energy products – Member States



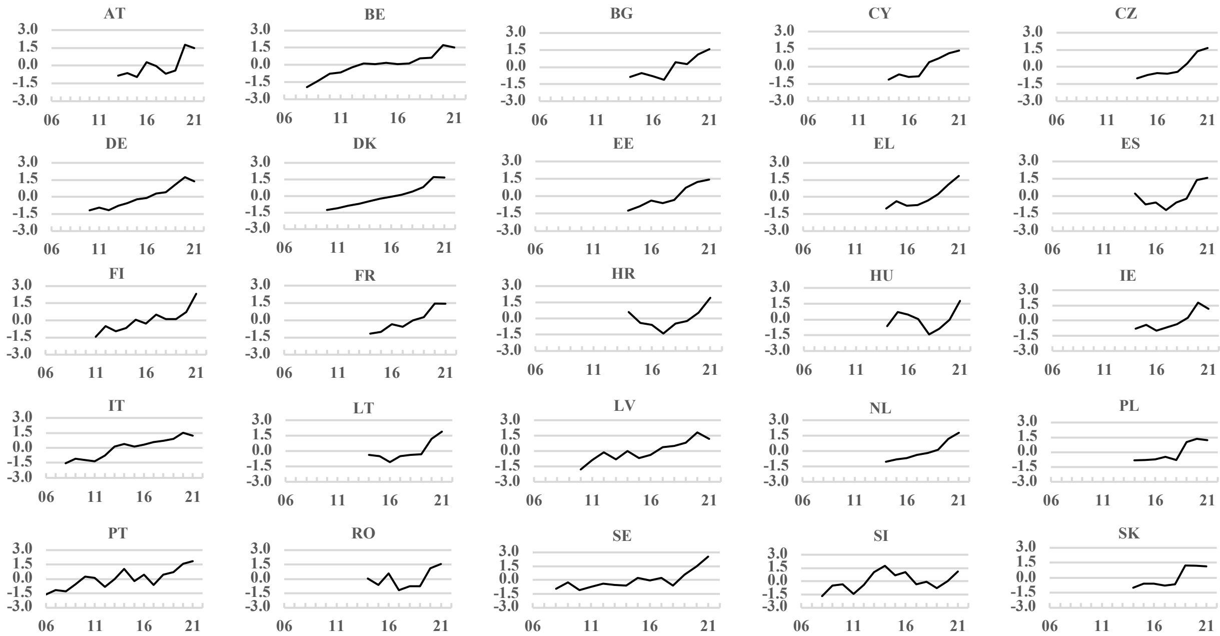
Source: Eurostat [env_ac_pefasu].
Note: Non-green products correspond to products [P08] to [P22]. Series are standardised to have zero mean and unit variance. The indicator displayed is expressed as 1 minus the share of non-green energy products. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1 for additional details.

Figure 16: Share of non-green energy products – Sectors



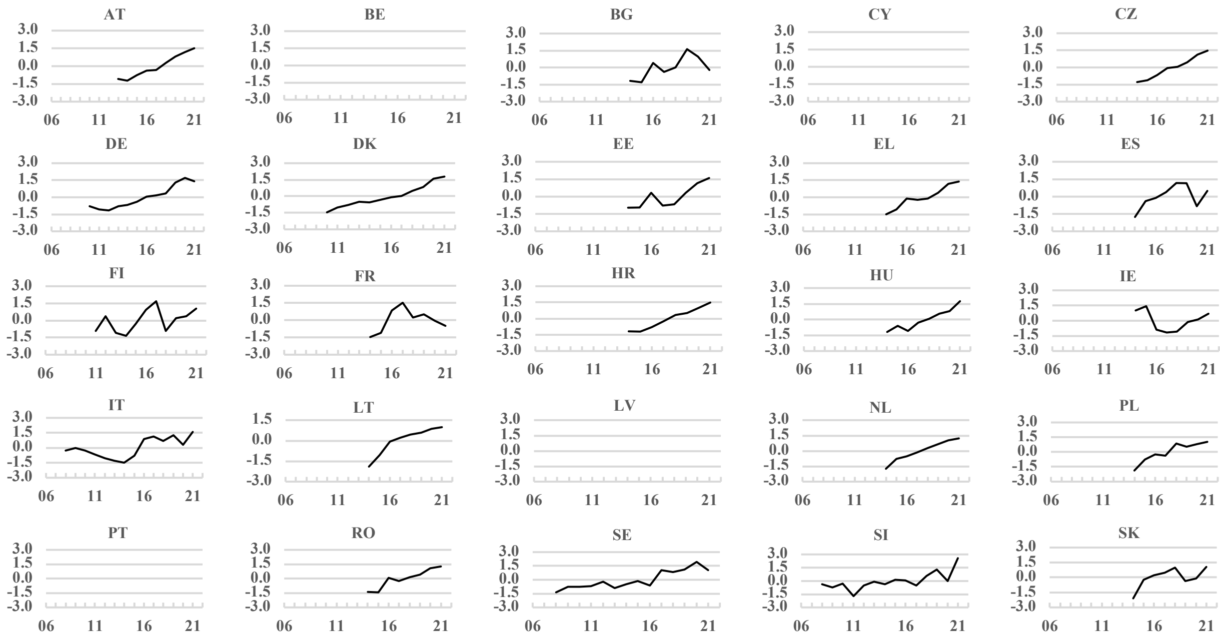
Source: Eurostat [env_ac_pefasu].
Note: Green products correspond to liquid biofuels [P24] and biogas [P25]. Series are standardised to have zero mean and unit variance. The indicator displayed is expressed as 1 minus the share of non-green energy products. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 17: Share of renewable natural resources – Member States



Source: Eurostat [env_ac_pefasu].
Note: Renewable natural resources correspond to resources [N03] to [N07]. Series are standardised to have zero mean and unit variance. See Section B1.1 and Table 1.

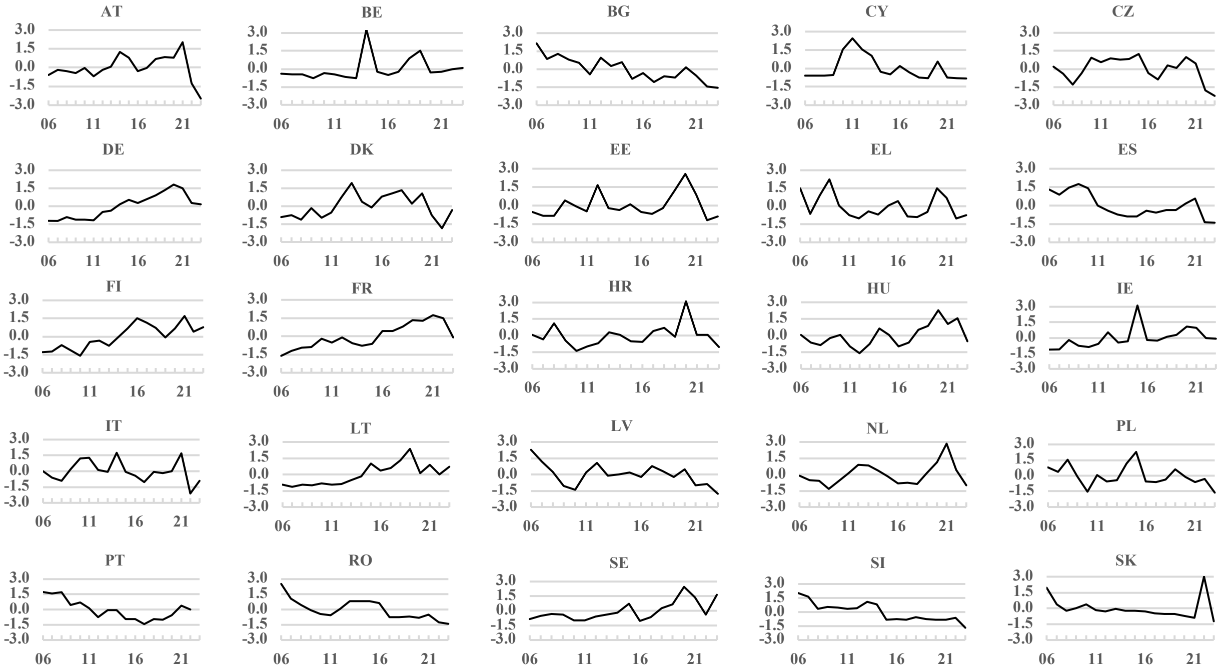
Figure 18: Share of non-renewable natural resources – Member States



Source: Eurostat [env_ac_pefasu].
Note: Non-renewable natural resources correspond to resources [N01] and [N02]. The indicator is expressed as 1 minus the share of non-renewable resources. BE, CY, LV and PT do not report using non-renewable natural resources (indicator is constant and equal to 1). Series are standardised to have zero mean and unit variance. See Section B1.1 and Table 1.

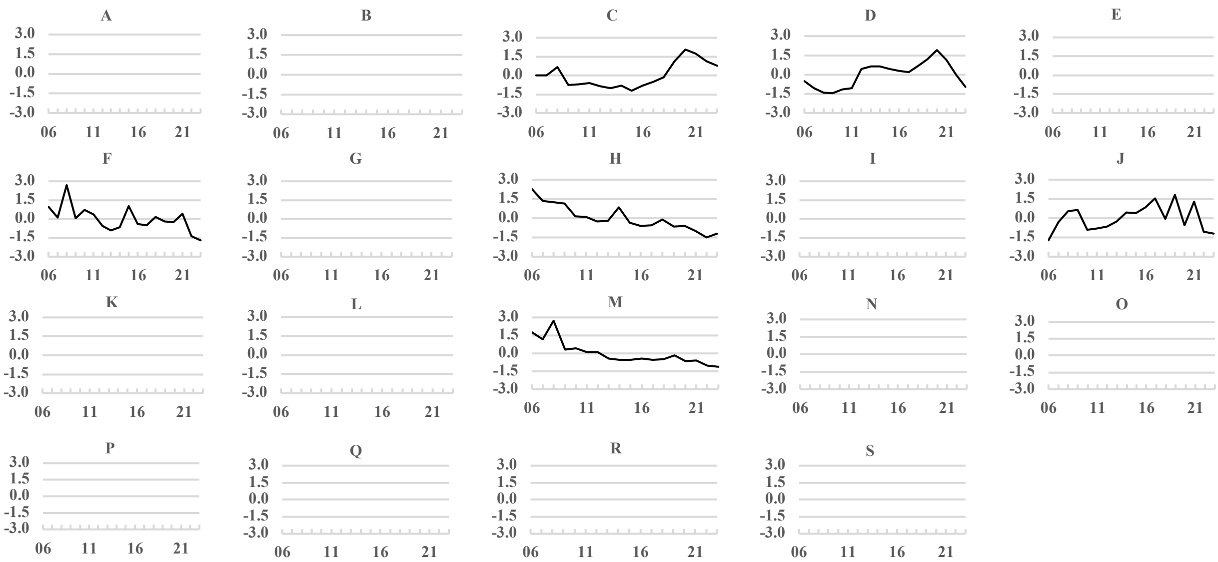
Capital

Figure 19: Investment in climate mitigation as a share of sectoral VA – Member States



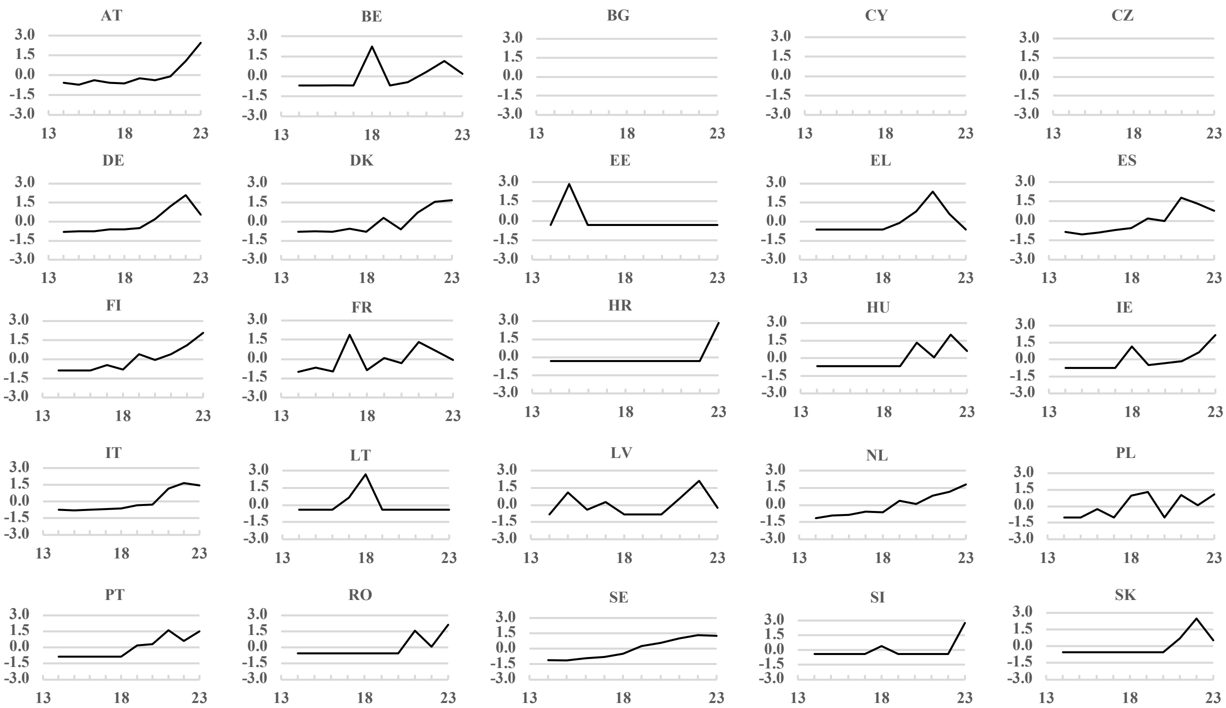
Source: Eurostat [env_ac_ccminv].
Note: Series are standardised to have zero mean and unit variance. Data for sector A and the service sector (net of H, J and M) is assumed to be 0. Data is not collected for sector B and is constant at 0 for sector E. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 20: Investment in climate mitigation as a share of sectoral VA – Sectors



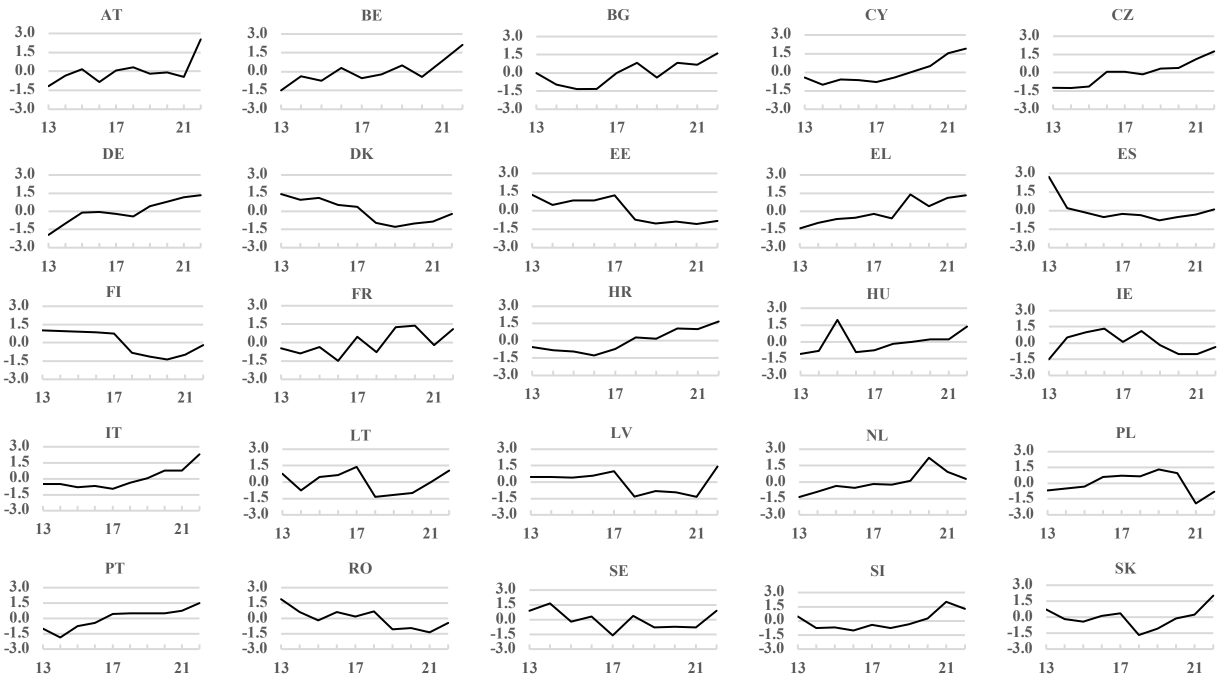
Source: Eurostat [env_ac_ccminv].
Note: Series are standardised to have zero mean and unit variance. Data for sector A and the service sector (G to S and net of H, J and M) is assumed to be 0. Data is not collected for sector B and is constant at 0 for sector E. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 21: Share of green bonds – Member States



Source: European Environment Agency.
Note: Green bonds as a share of total bonds issued . BG, CY and CZ do not report issuance of green bonds (indicator is constant at 0). Series are standardised to have zero mean and unit variance. See Section B1.1 and Table 1.

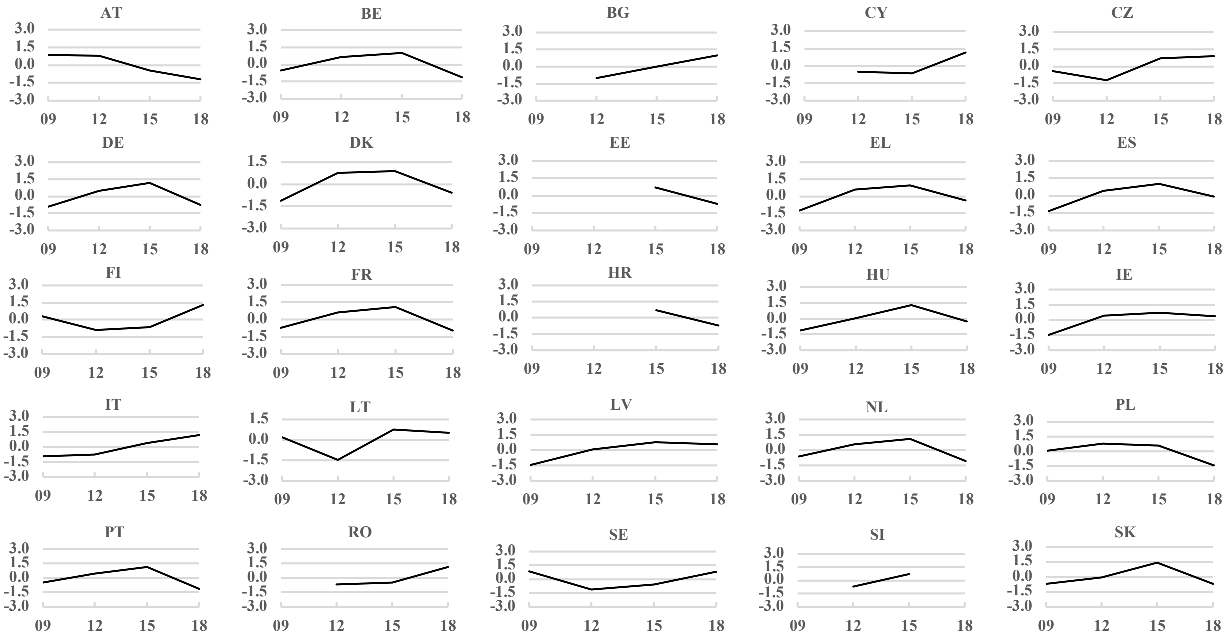
Figure 22: Eco-innovation input index – Member States



Source: European Environment Agency.
Note: Series are standardised to have zero mean and unit variance. See Section B1.1 and Table 1.

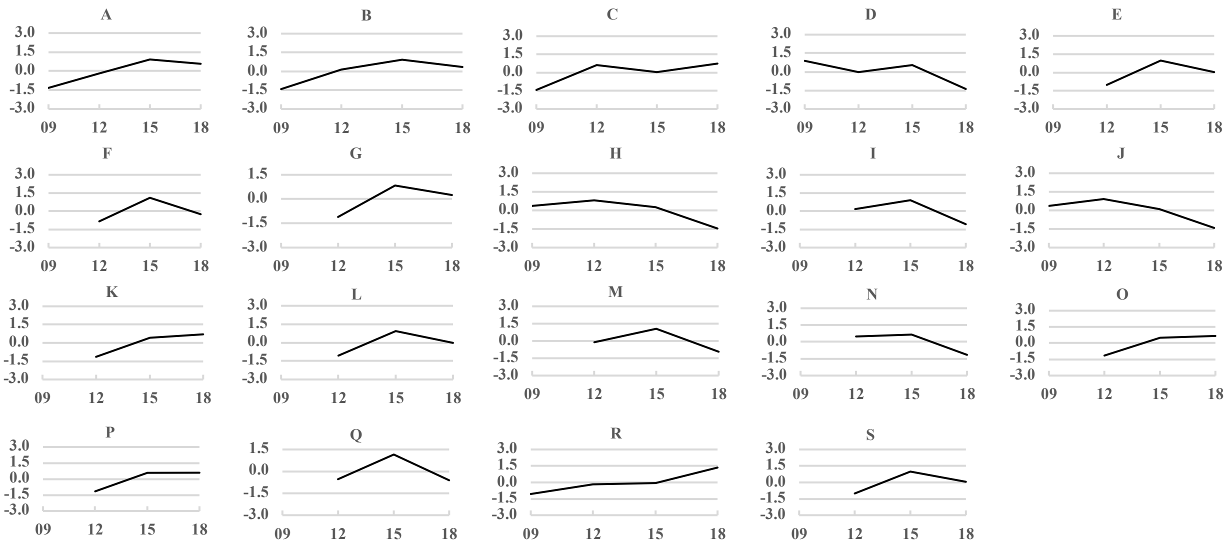
Land use

Figure 23: Land use as a share of total land – Member States



Source: Eurostat [lan_use_ovw].
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 24: Land use as a share of total land – Sectors

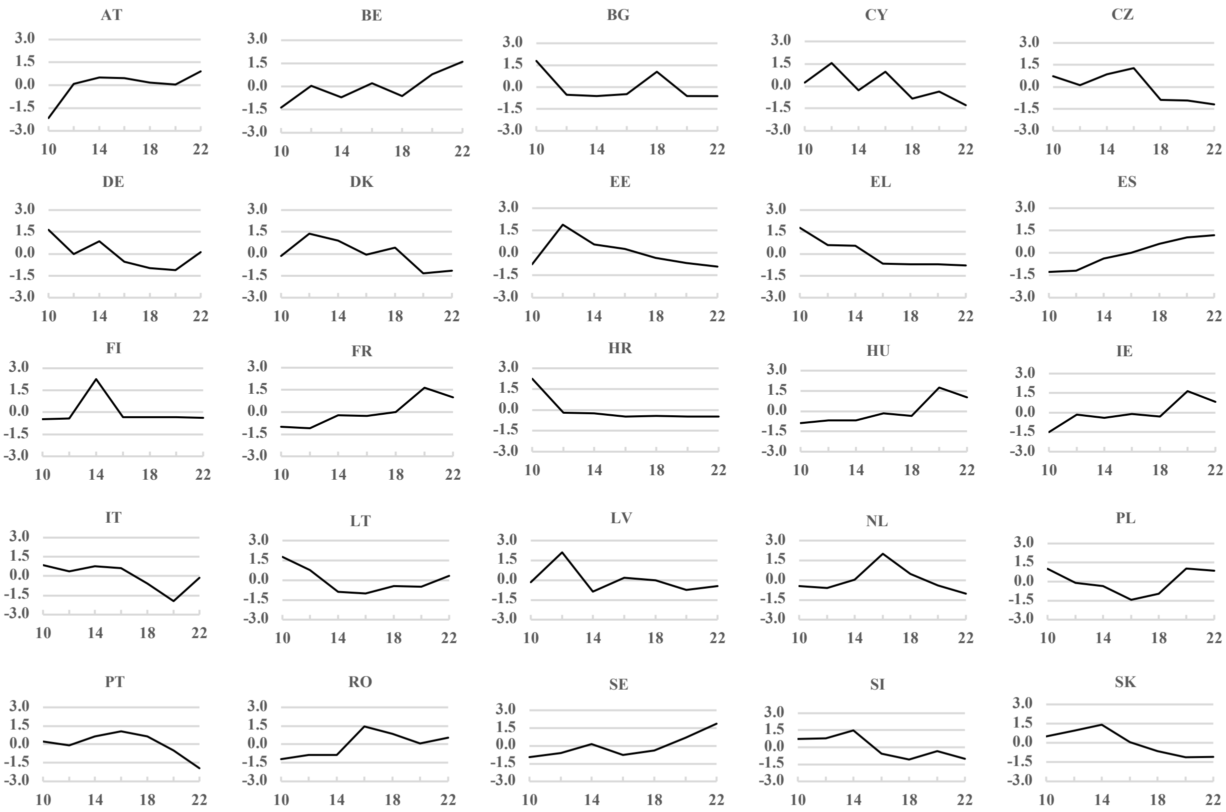


Source: Eurostat [lan_use_ovw].
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

B2.2. *Sector – Process*

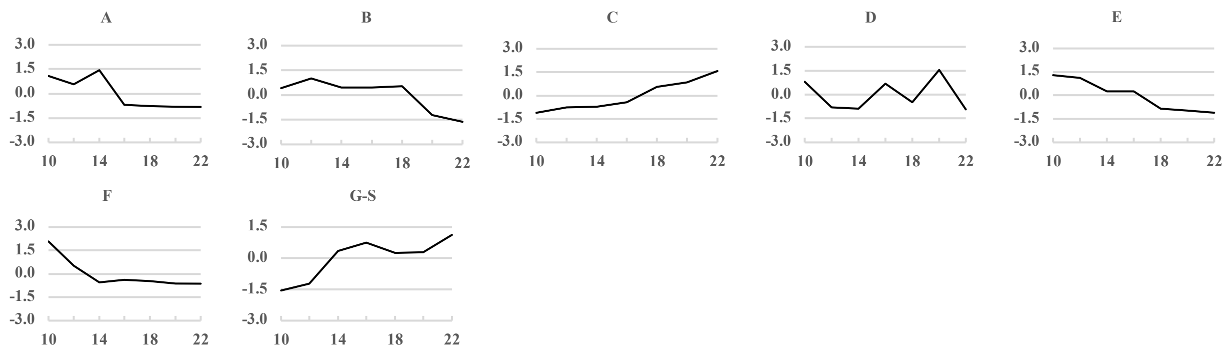
Waste and circularity

Figure 25: VA per 1 tonne of waste – Member States



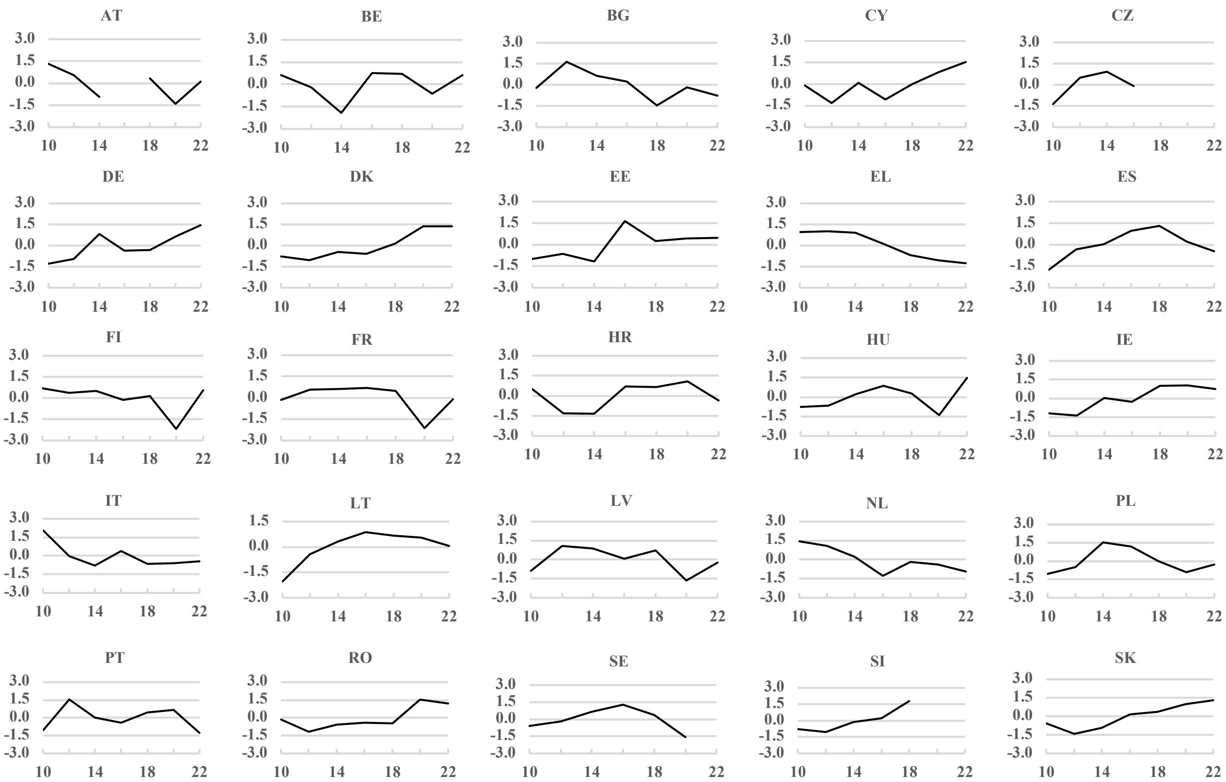
Source: Eurostat [env_wasgen].
Note: Data on waste is collected every two years. Series are standardised to have zero mean and unit variance. The series at MS level are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1 for additional details.

Figure 26: VA per 1 tonne of waste – Sectors



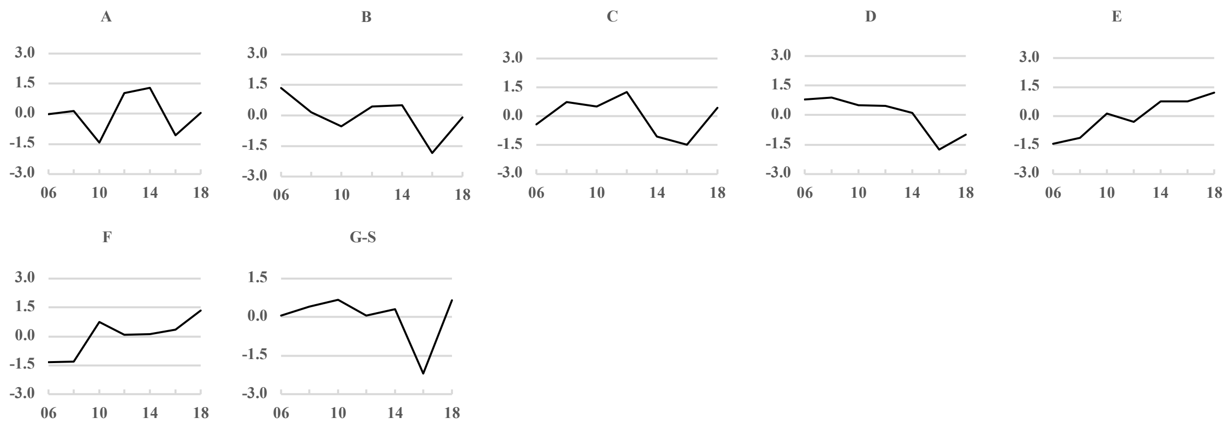
Source: Eurostat [env_wasgen].
Note: Data on waste is collected every two years. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 27: Share of chemical waste – Member States



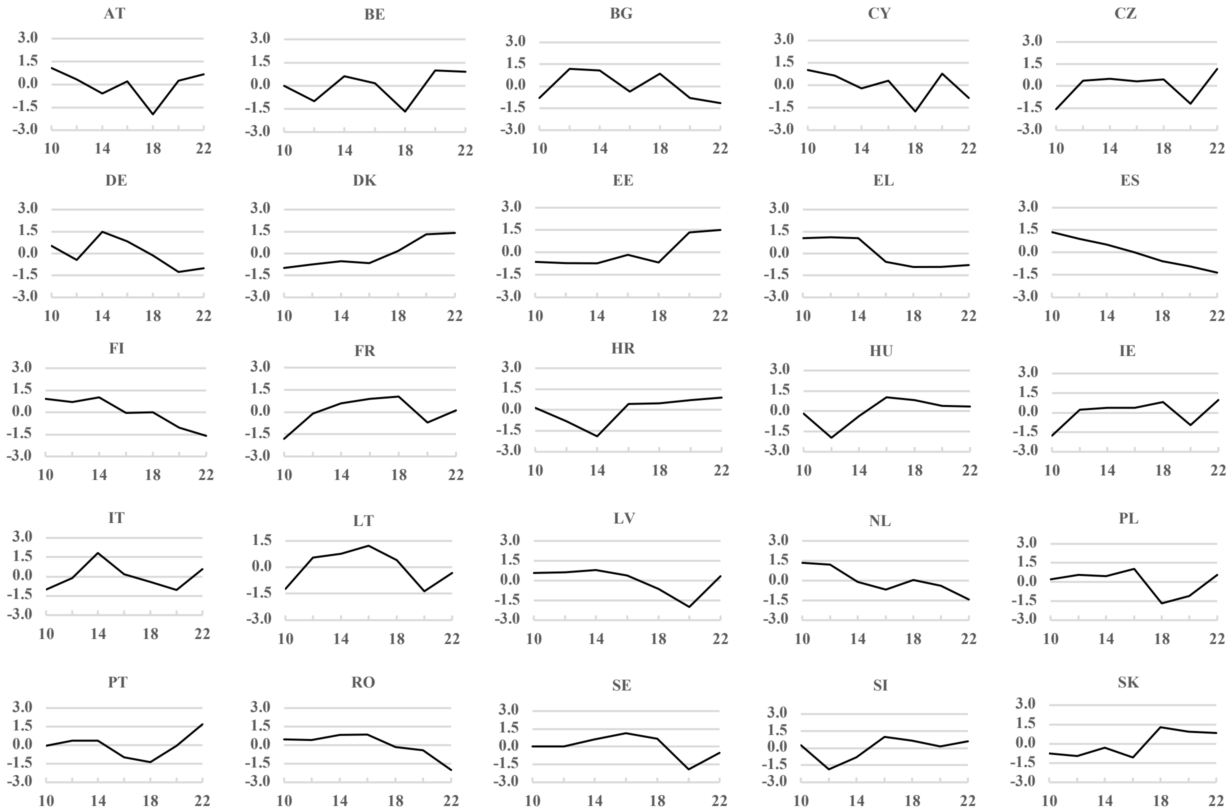
Source: Eurostat [env_wasgen].
Note: Chemical wastes include categories spent solvents [W011] to health care and biological wastes [W05]. The indicator is expressed as 1 minus the share of chemical wastes. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 28: Share of chemical waste – Sectors



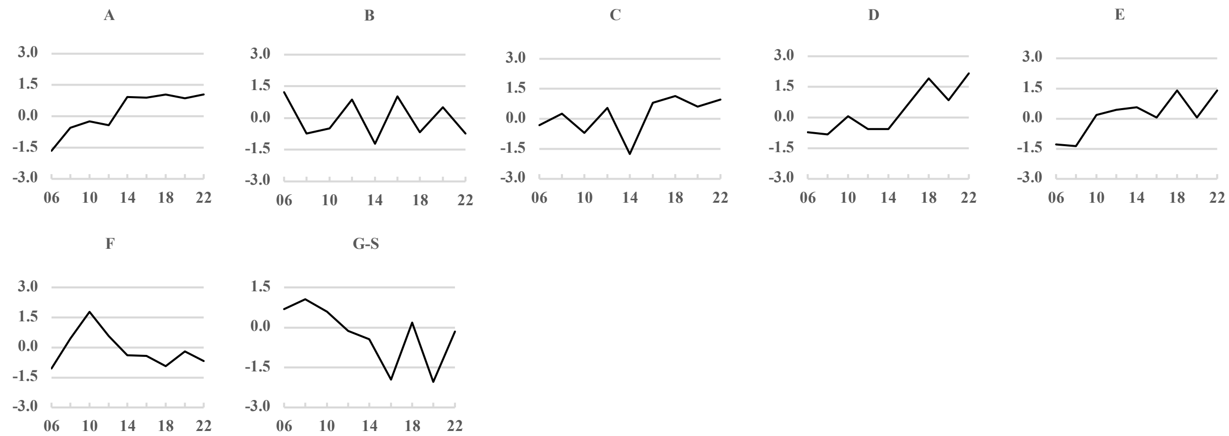
Source: Eurostat [env_wasgen].
Note: Chemical wastes include categories spent solvents [W011] to health care and biological wastes [W05]. The indicator is expressed as 1 minus the share of chemical wastes. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 29: Share of hazardous waste – Member States



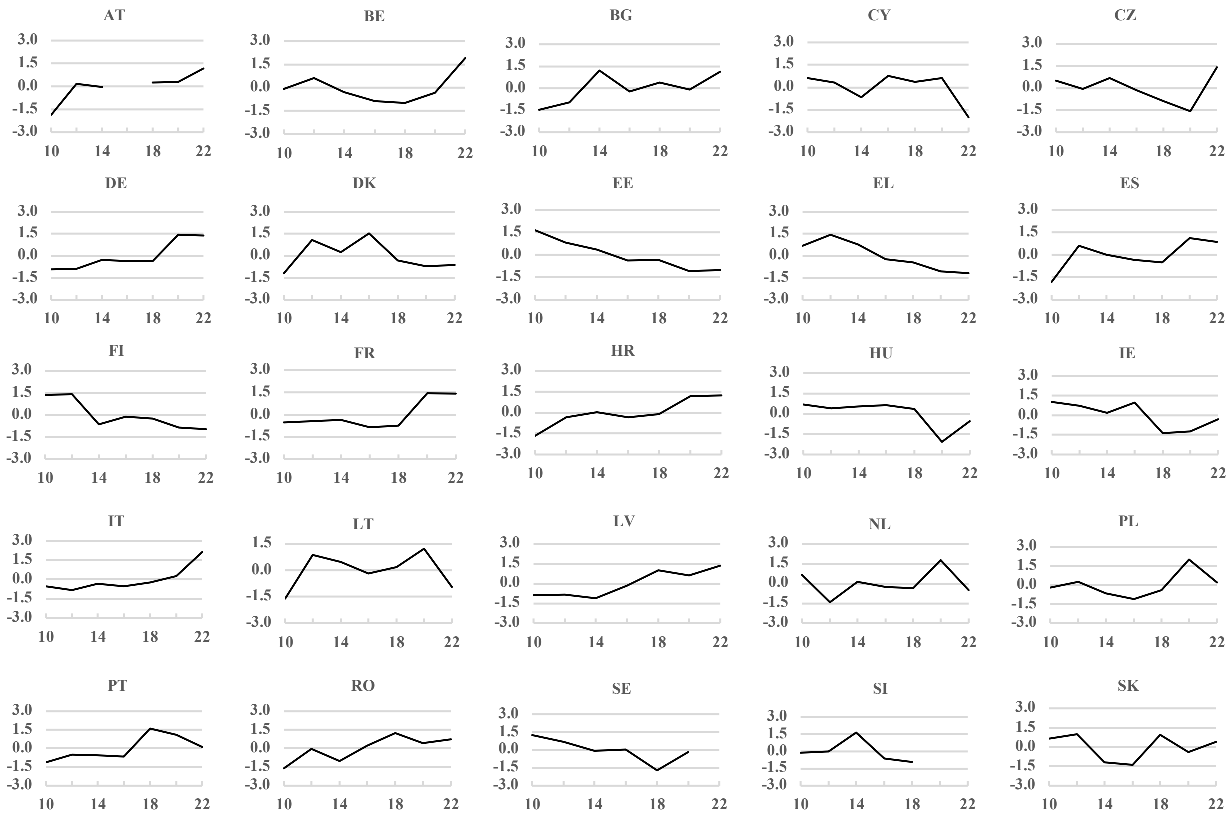
Source: Eurostat [env_wasgen].
Note: The hazardous nature is determined according to the concentration of specific substances in the waste. The indicator is expressed as 1 minus the share of hazardous wastes. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 30: Share of hazardous waste – Sectors



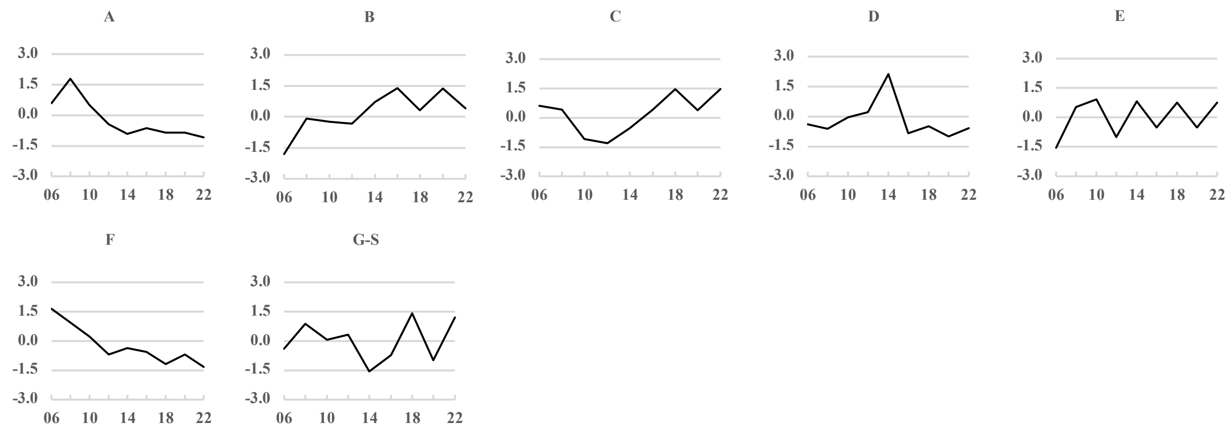
Source: Eurostat [env_wasgen].
Note: The hazardous nature is determined according to the concentration of specific substances in the waste. The indicator is expressed as 1 minus the share of hazardous wastes. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 31: Share of recyclable waste – Member States



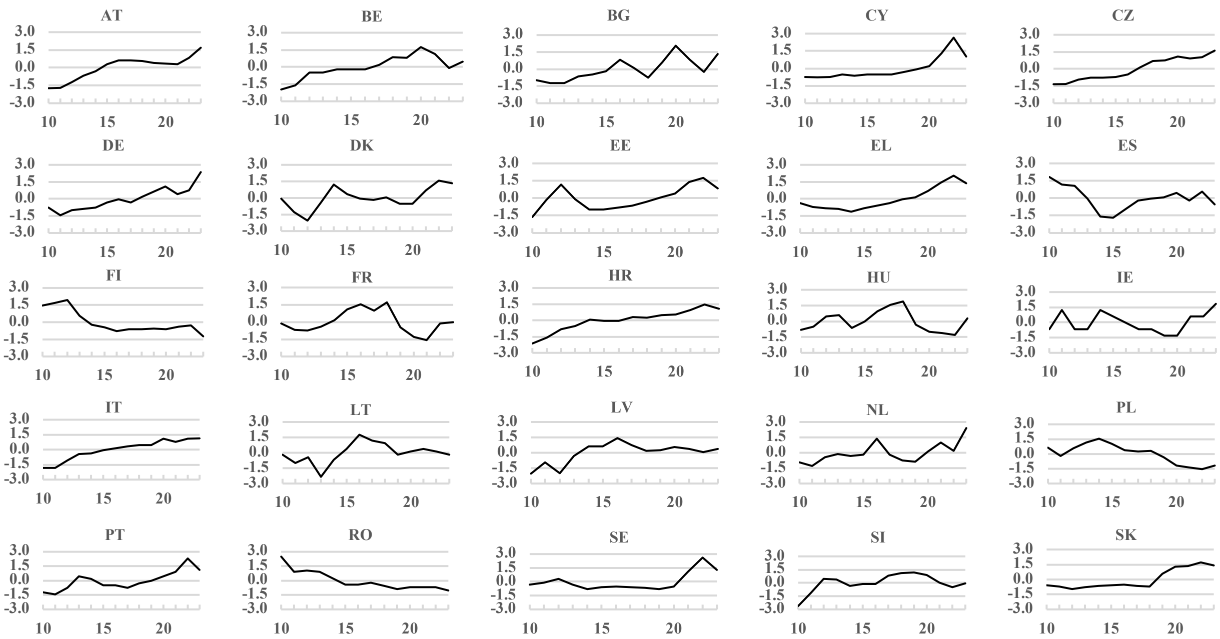
Source: Eurostat [env_wasgen].
Note: Recyclable wastes include categories metallic wastes, ferrous [W061] to textile wastes [W076]. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 32: Share of recyclable waste – Sectors



Source: Eurostat [env_wasgen].
Note: Recyclable wastes include categories metallic wastes, ferrous [W061] to textile wastes [W076]. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

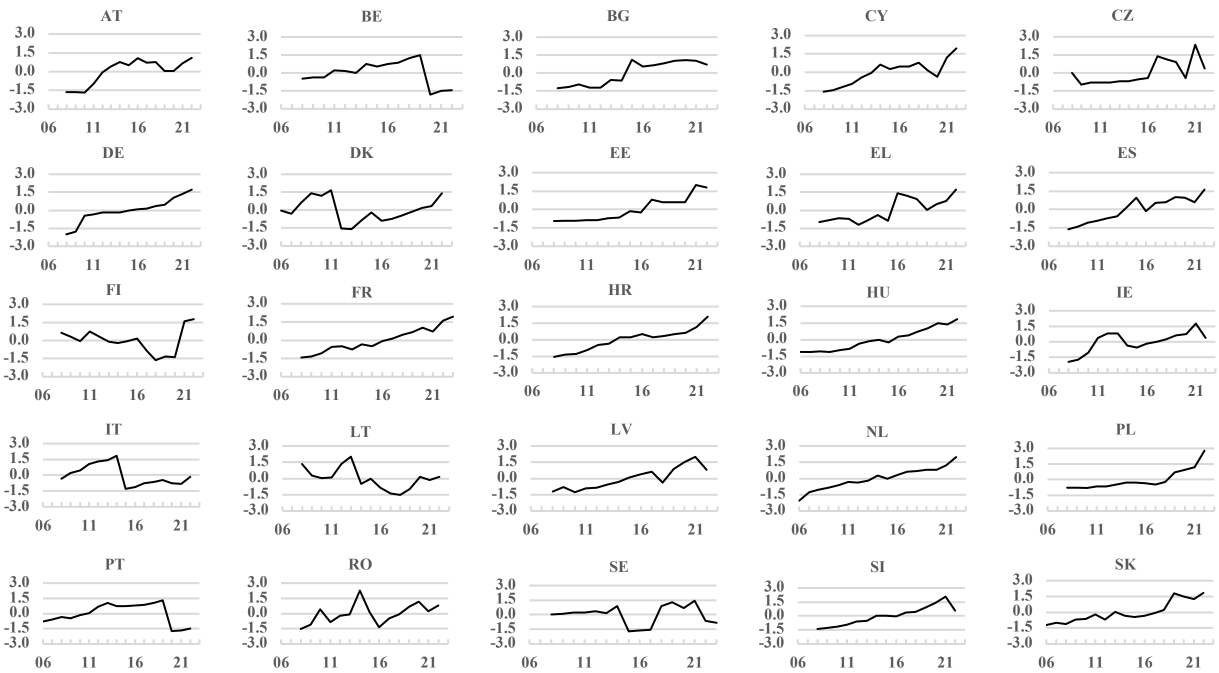
Figure 33: Circularity material use rate – Member States



Source: Eurostat [env_ac_cur].
Note: Series are standardised to have zero mean and unit variance. See Section B1.1 and Table 1.

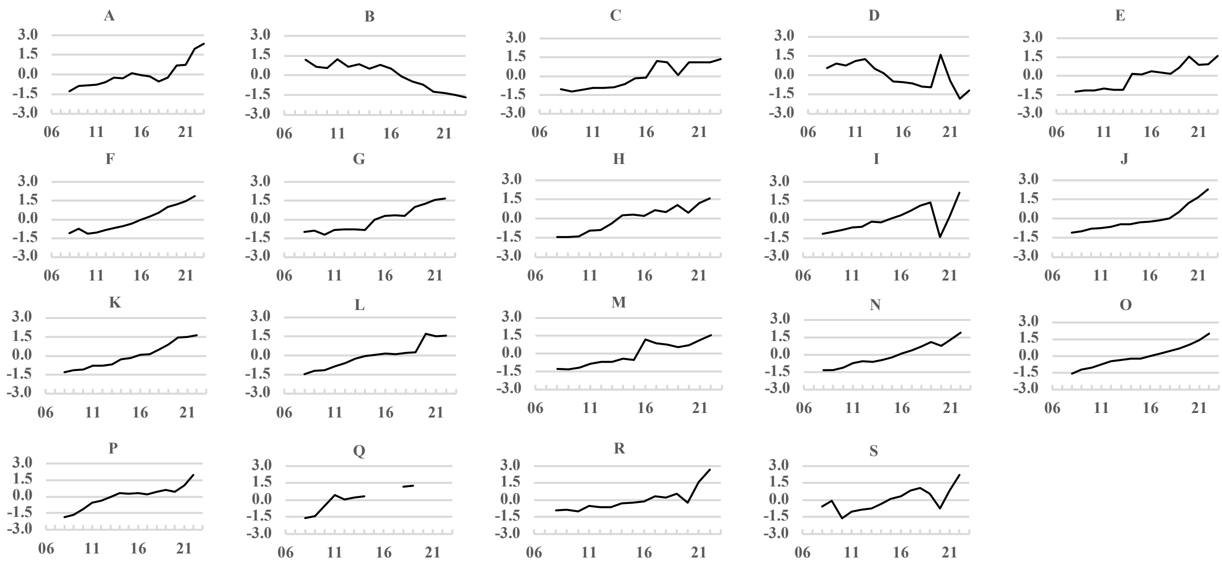
Air emissions

Figure 34: VA per 1 tonne of ACG emission – Member States



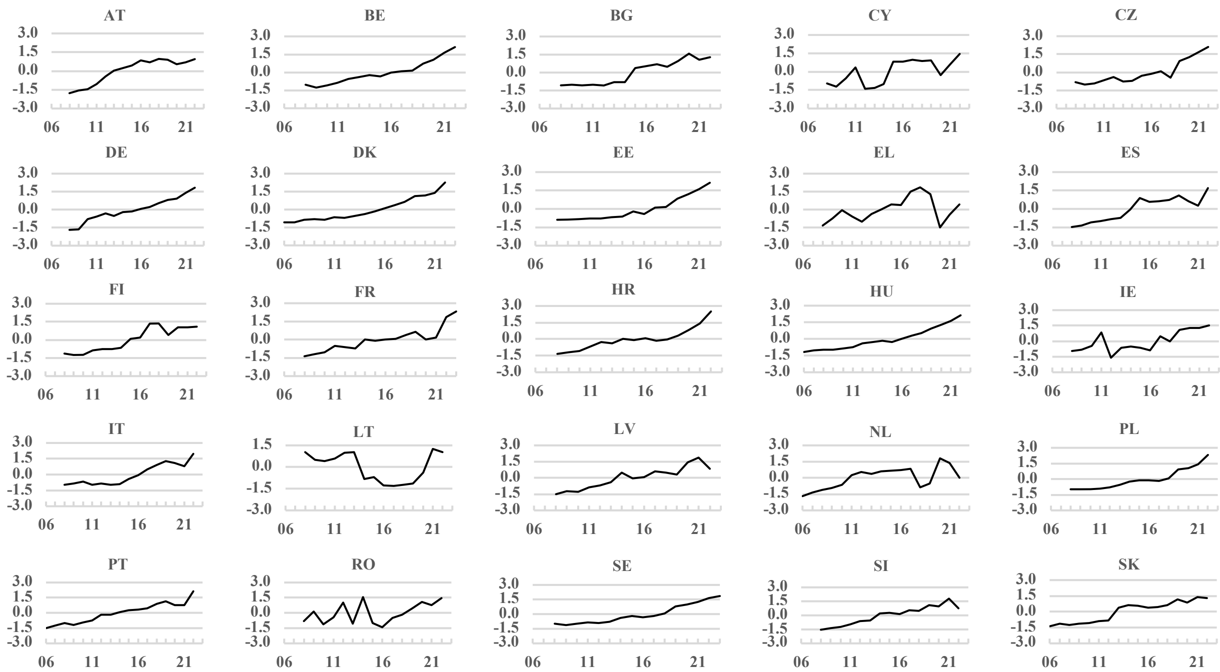
Source: Eurostat [env_ac_ainah_r2].
Note: Acidifying gas corresponds to Sulphur oxides,[SOx], Nitrogen oxides [N0x] and Ammonia [NH3]. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 35: VA per 1 tonne of ACG emission – Sectors



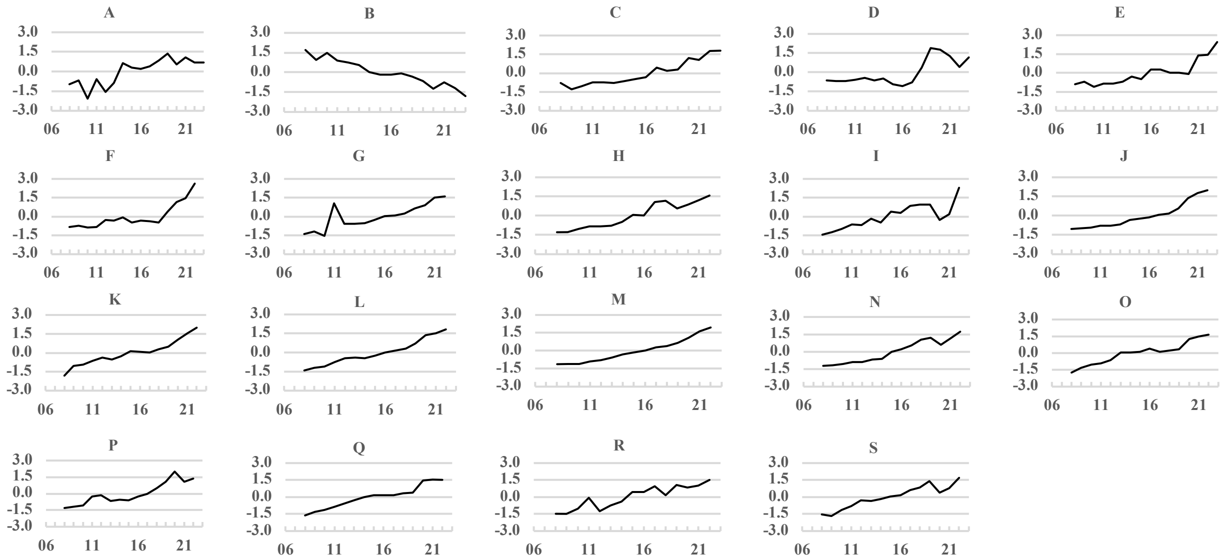
Source: Eurostat [env_ac_ainah_r2].
Note: Acidifying gas corresponds to Sulphur oxides,[SOx], Nitrogen oxides [N0x] and Ammonia [NH3]. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 36: VA per 1 tonne of GHG emission – Member States



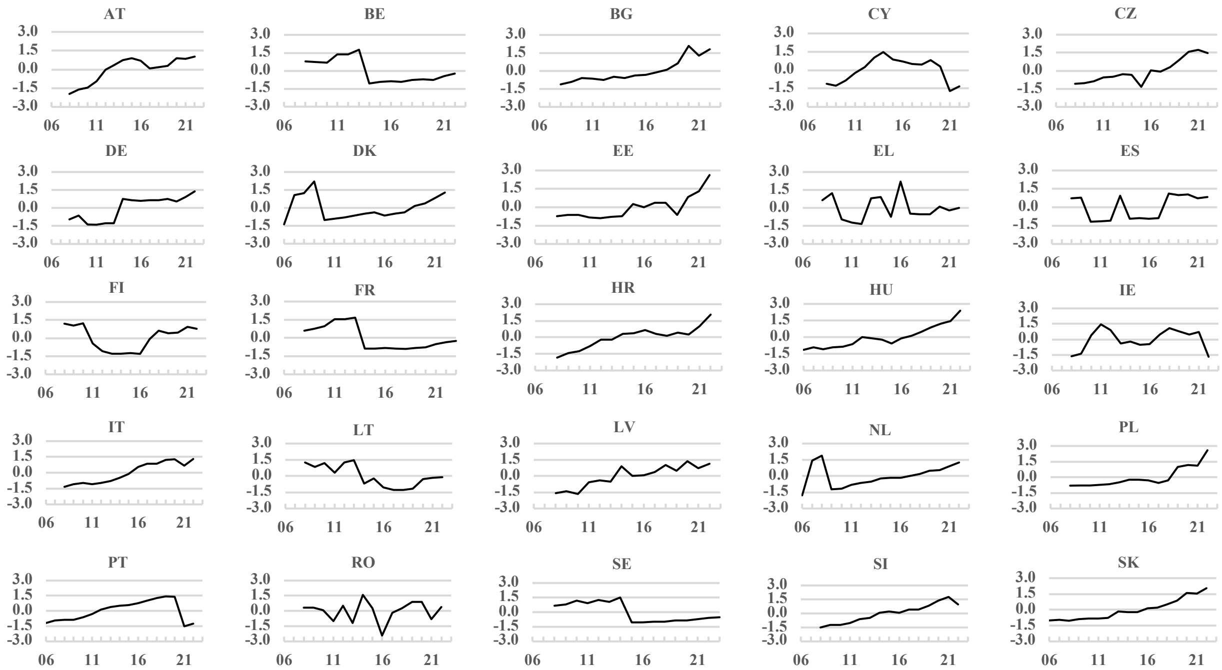
Source: Eurostat [env_ac_ainah_r2].
Note: Greenhouse gases include Carbon dioxide [CO2], Nitrous oxide [N20], Methane [CH4], Perfluorocarbons [PFC], Hydrofluorocarbons [HFC] and Sulphur hexafluoride [SF6]. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 37: VA per 1 tonne of GHG emission – Sectors



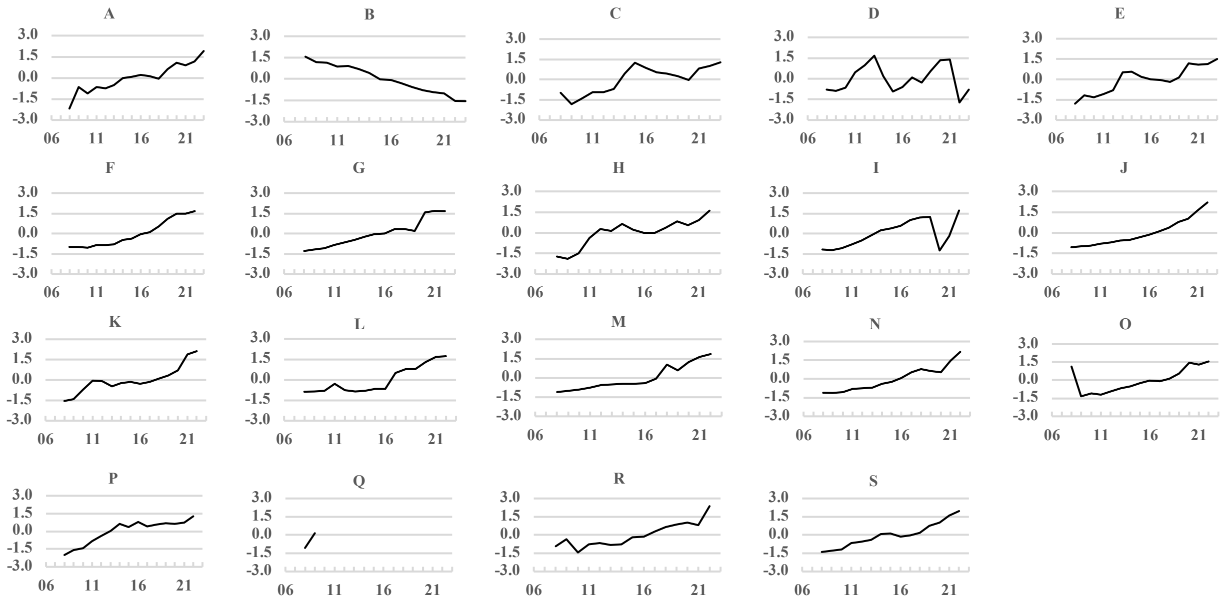
Source: Eurostat [env_ac_ainah_r2].
Note: Greenhouse gases include Carbon dioxide [CO2], Nitrous oxide [N20], Methane [CH4], Perfluorocarbons [PFC], Hydrofluorocarbons [HFC] and Sulphur hexafluoride [SF6]. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 38: VA per 1 tonne of O3PR emission – Member States



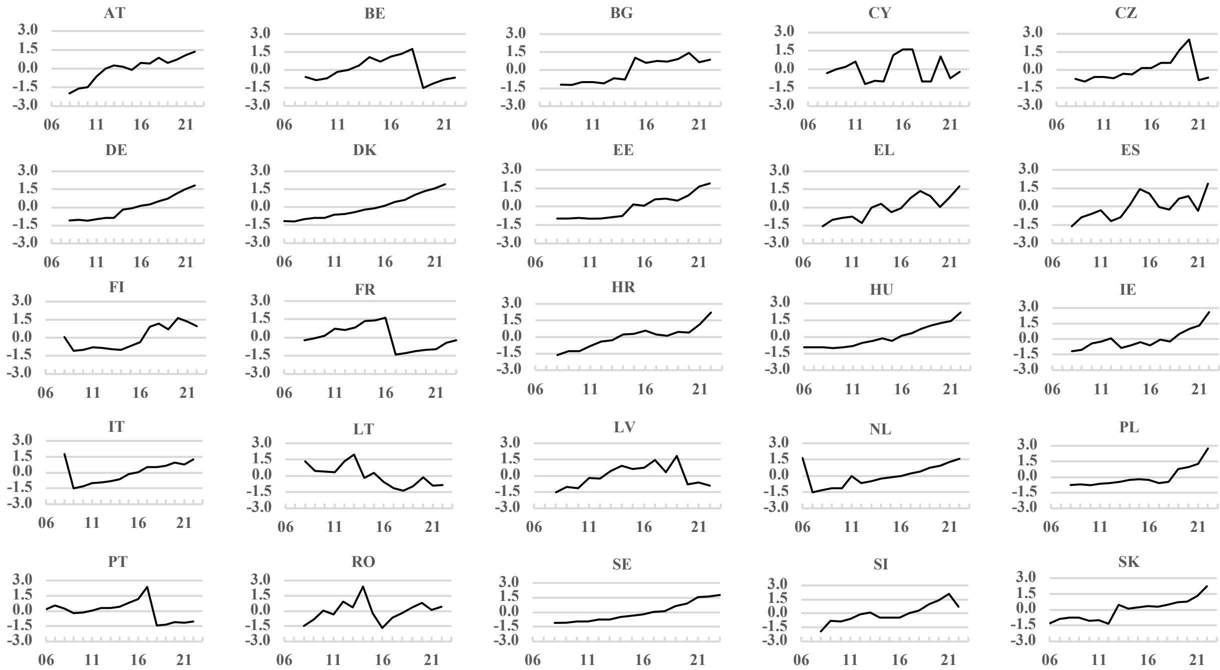
Source: Eurostat [env_ac_ainah_r2].
Note: Tropospheric ozone precursors include Non-methane volatile organic compounds [NMVOC], Nitrogen oxides [NOx], Carbon monoxide [CO] and Methane [CH4]. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 39: VA per 1 tonne of O3PR emission – Sectors



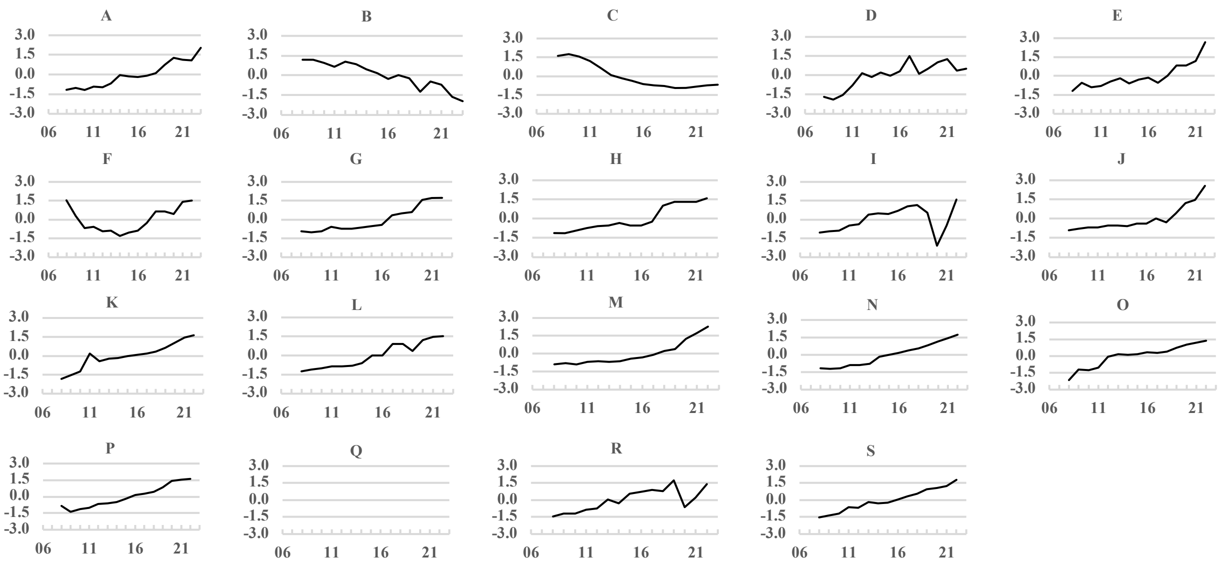
Source: Eurostat [env_ac_ainah_r2].
Note: Tropospheric ozone precursors include Non-methane volatile organic compounds [NMVOC], Nitrogen oxides [NOx], Carbon monoxide [CO] and Methane [CH4]. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 40: VA per 1 tonne of particulate matter emission – Member States



Source: Eurostat [env_ac_ainah_r2].
Note: Particulate matter smaller than 2.5 microns. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

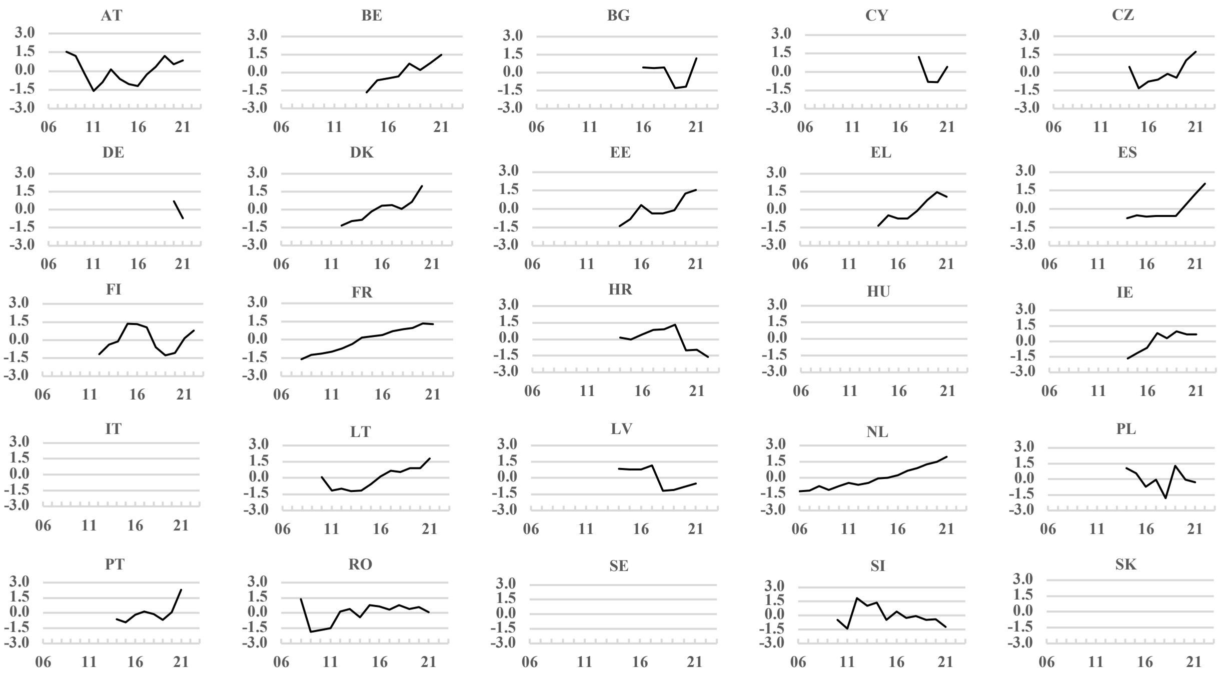
Figure 41: VA per 1 tonne of particulate matter emission – Sectors



Source: Eurostat [env_ac_ainah_r2].
Note: Particulate matter smaller than 2.5 microns. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

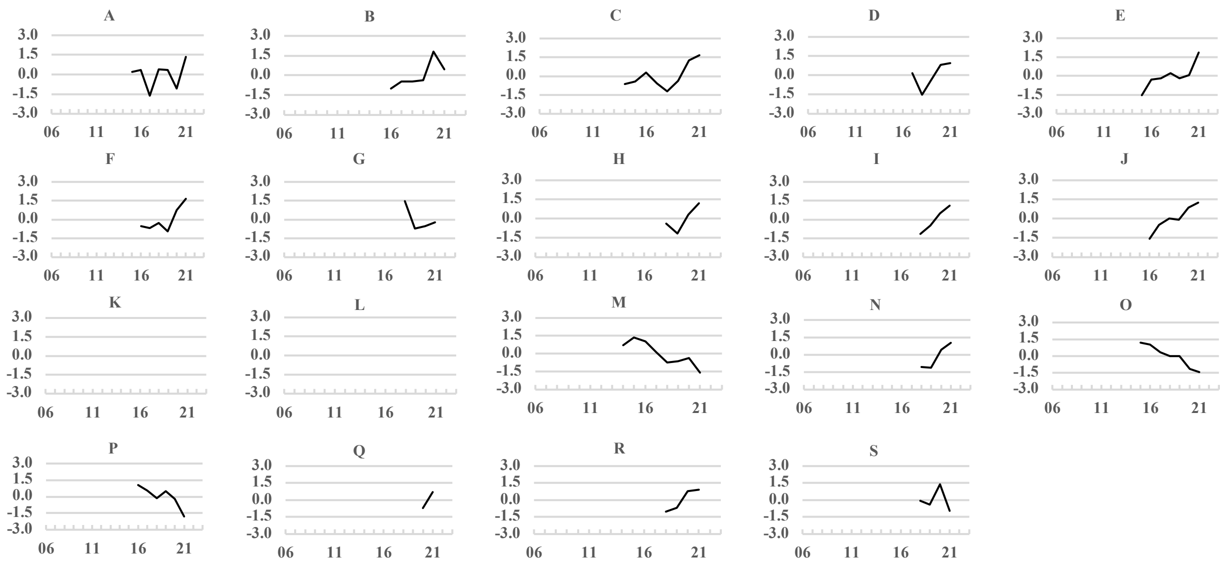
B2.3. Sector – Output

Figure 42: EGSS as a share of VA – Member States



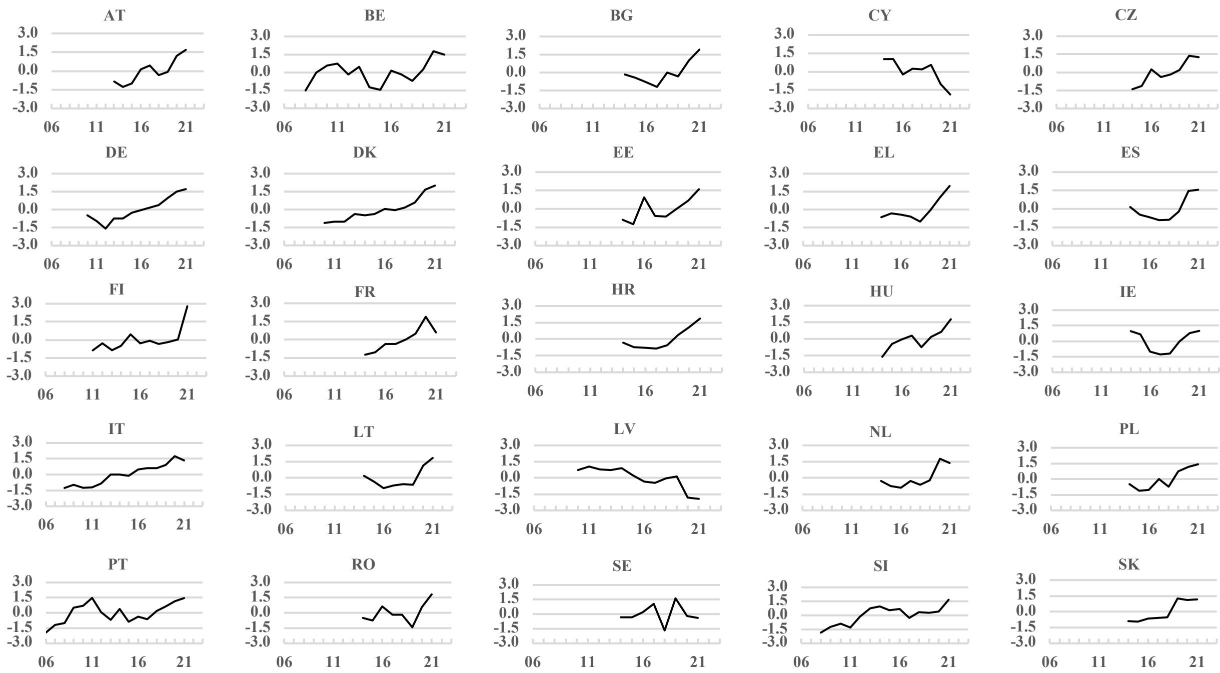
Source: Eurostat [env_ac_ainah_r2].
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 43: EGSS as a share of VA – Sectors



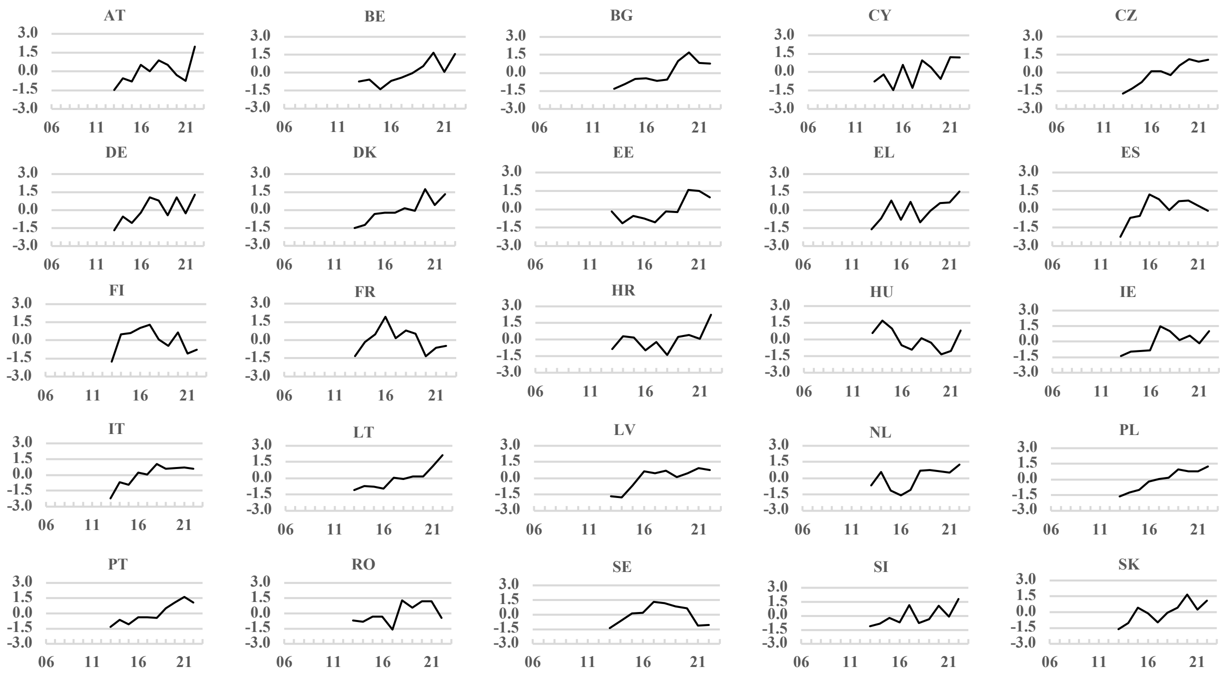
Source: Eurostat [env_ac_ainah_r2].
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. The series at sectoral level are obtained by averaging series across MS. Averages are displayed if data is available for at least 20 MS (out of the 25). See Section B1.1 and Table 1 for additional details and https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

Figure 44: Supply of nongreen energy products as a share of total energy supplied – Member States



Source: Eurostat [env_ac_ainah_r2].
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

Figure 45: Eco-innovation output index – Member States

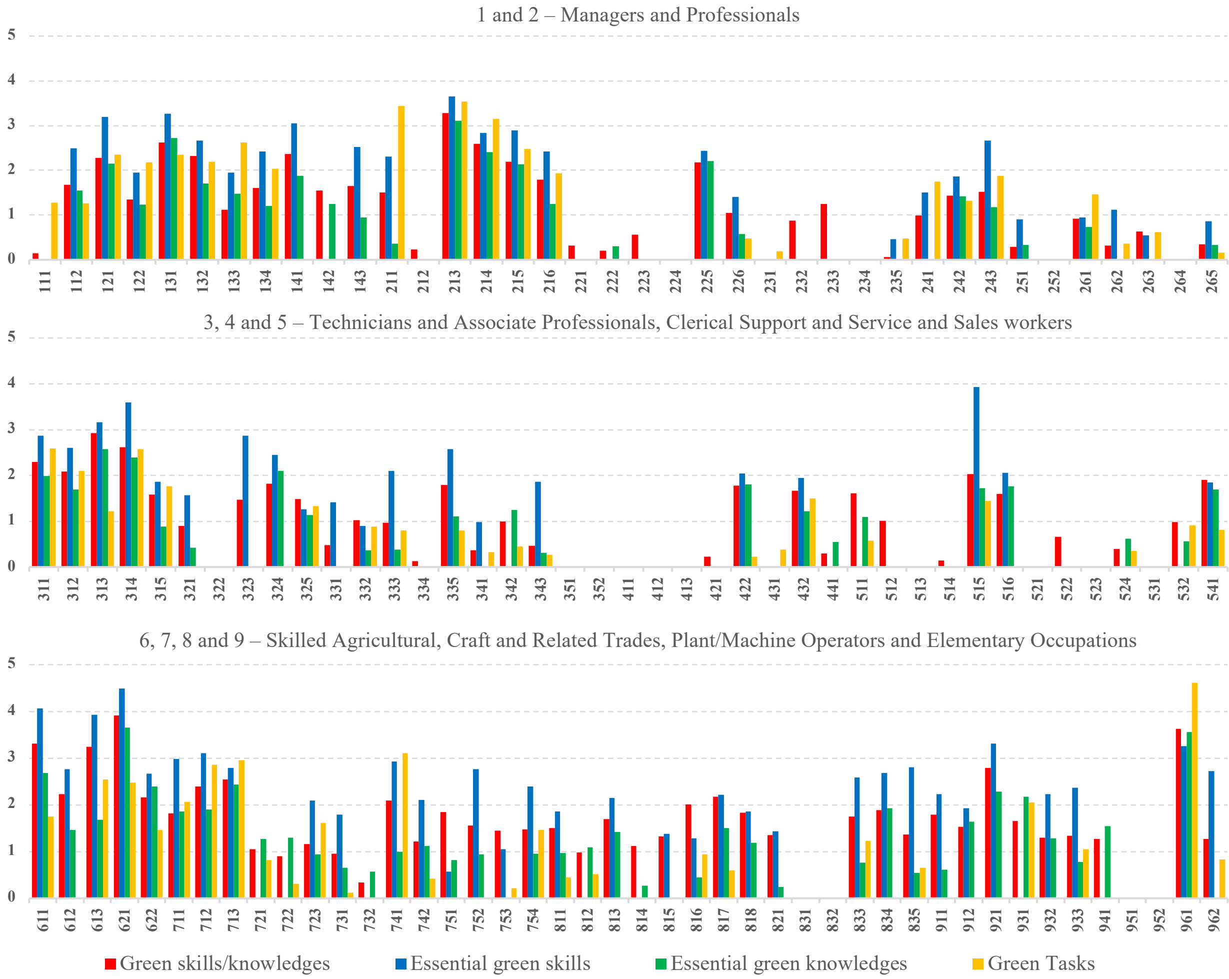


Source: European environmental agency.
Note: The indicator is expressed as 1 minus the share of land used. Series are standardised to have zero mean and unit variance. Number by MS are obtained by averaging series across sectors. Averages are displayed if data is available for at least 15 sectors (out of the 19). See Section B1.1 and Table 1.

B2.4. Occupation – Input and Process

Skills and tasks

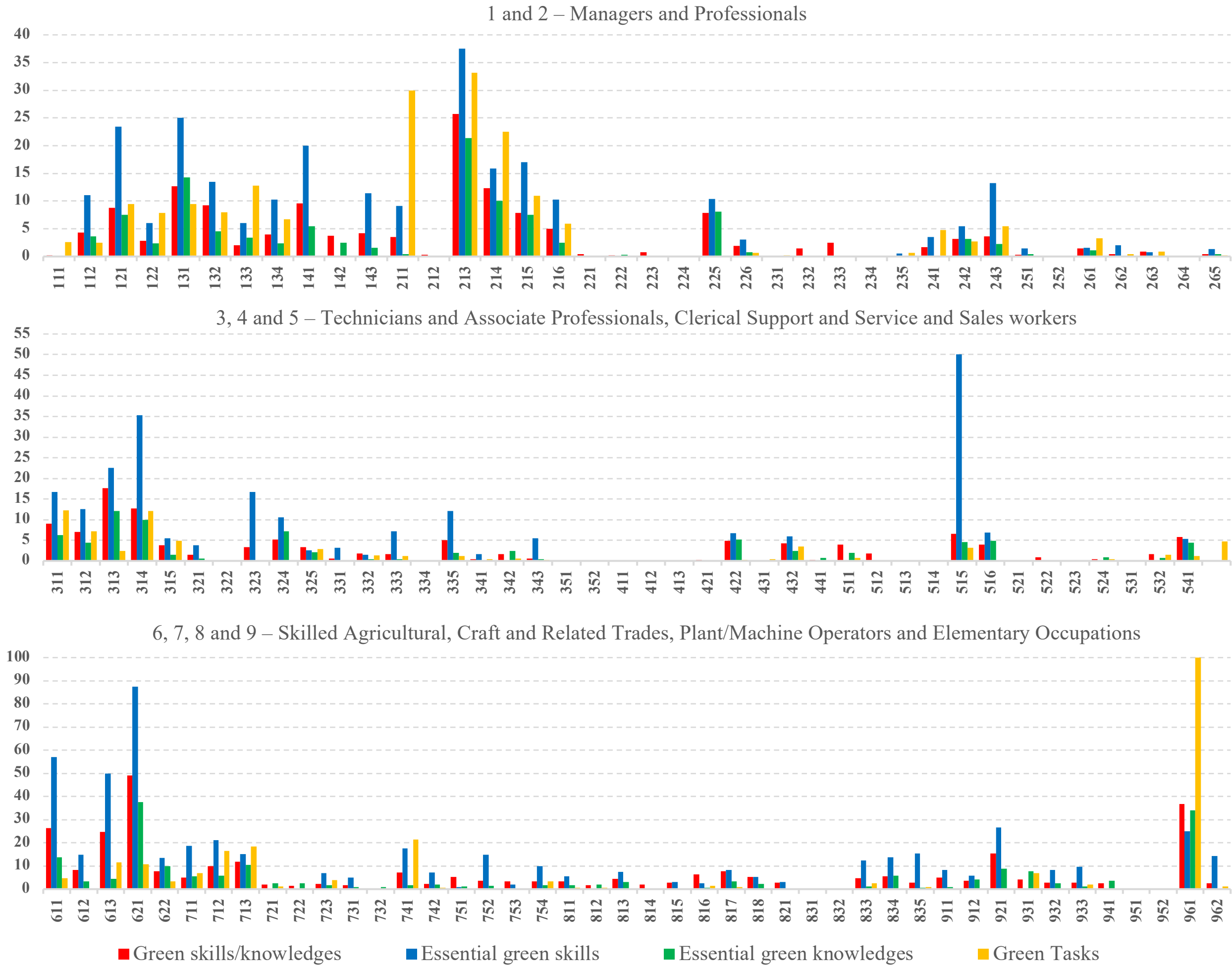
Figure 46: Green skills and tasks



Source: ESCO and O*NET

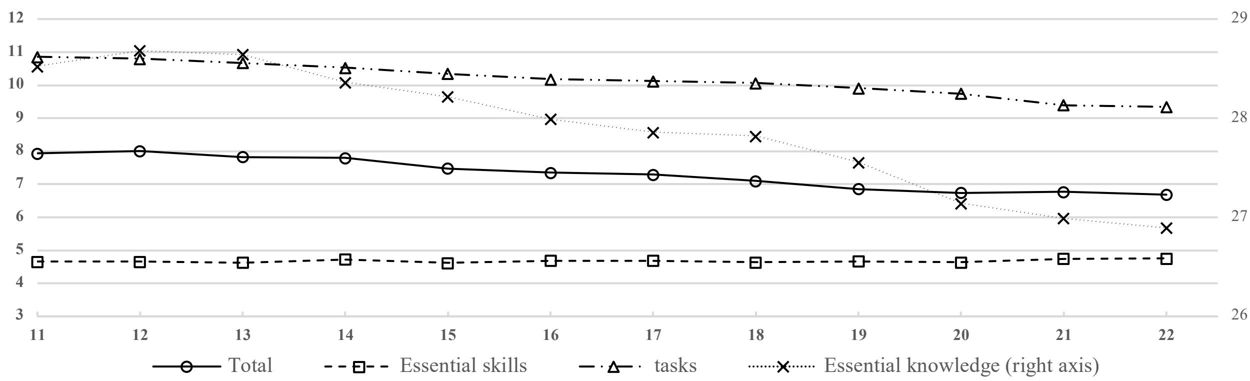
Note: The indicator is expressed as ln (1 + the share of the relevant green concepts). The full list of 3 digit ISCO occupations by major groups can be obtained from the following link https://ilostat.ilo.org/classification-occupation/. See Section B1.2 and B1.4 for additional details on the indicators.

Figure 47: Green skills and tasks – untransformed indicators



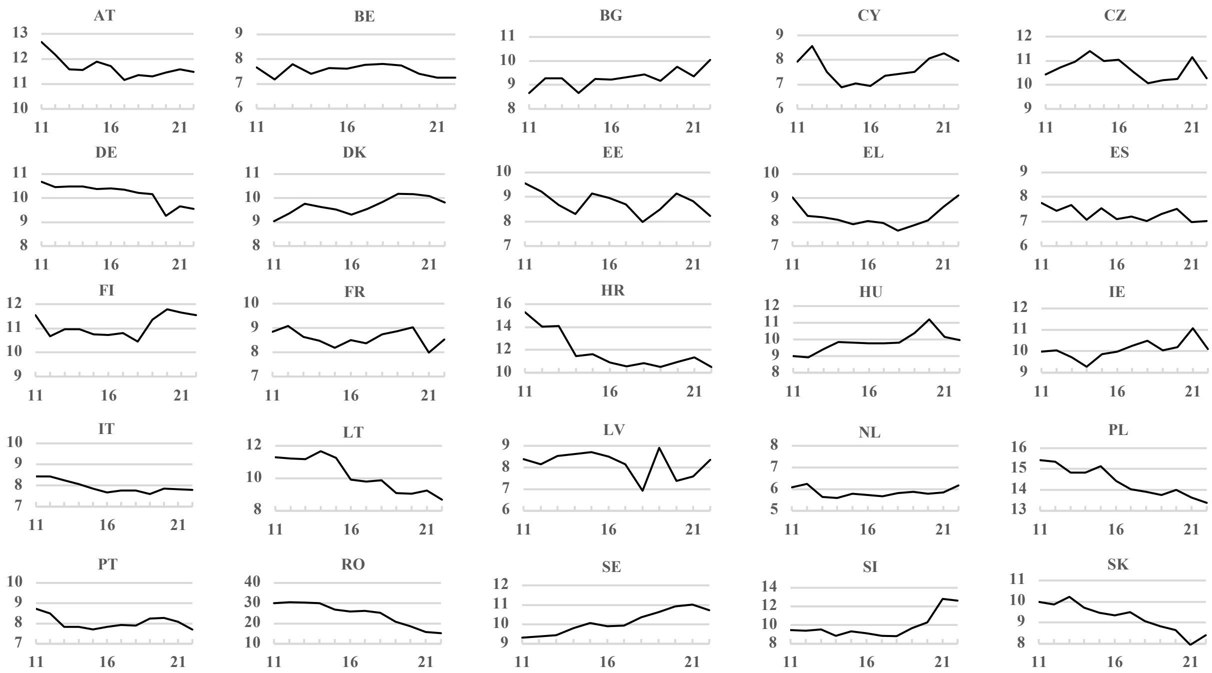
Source: ESCO and O*NET
Note: Share of the relevant green concepts. The full list of 3 digit ISCO occupations by major groups can be obtained from the following link: https://ilostat.ilo.org/classification-occupation/. See Section B1.2 and B1.4 for additional details on the indicators.

Figure 48: Share of green jobs in the EU27 – skills, knowledge and tasks



Source: ESCO, O*NET and EU-LFS.
Note: Share of green jobs in percentages according to the skills and task indicators. An occupation is considered green if the share of the relevant green concepts is greater than 10%. The share of green jobs obtained from essential knowledge is displayed on the right axis. Data for BG and SI, available only at 2-digit ISCO level, is not included in the computation.

Figure 49: Share of green jobs – tasks



Source: O*NET and EU-LFS.
Note: Share of green jobs in percentages according to the skills and task indicators. An occupation is considered green if the share of the relevant green concepts is greater than 10%. The share of green jobs obtained from essential knowledge is displayed on the right axis. Data for BG and SI, available only at 2-digit ISCO level, is not included in the computation.

B3. Multivariate statistical analysis

Given the very large number of series, the tools used for step 4 of Nardo et al. (2005) are not very appropriate to the current situation. For instance, it is very difficult to display correlations matrices or Cronbach alphas for more than 4000 series. Likewise, it is virtually impossible to expect one latent component extracted from PCA to explain a significant share of the variance of such a large number of series. It would be possible to perform a 'pooled' PCA analysis at the indicator level (i.e. ignoring the MS and sector dimensions) but it is unclear whether these results would be relevant given the DFM exploit the variations along these dimensions.

Nevertheless, this section will eventually present the PCA results computed to initialise the EM algorithm.

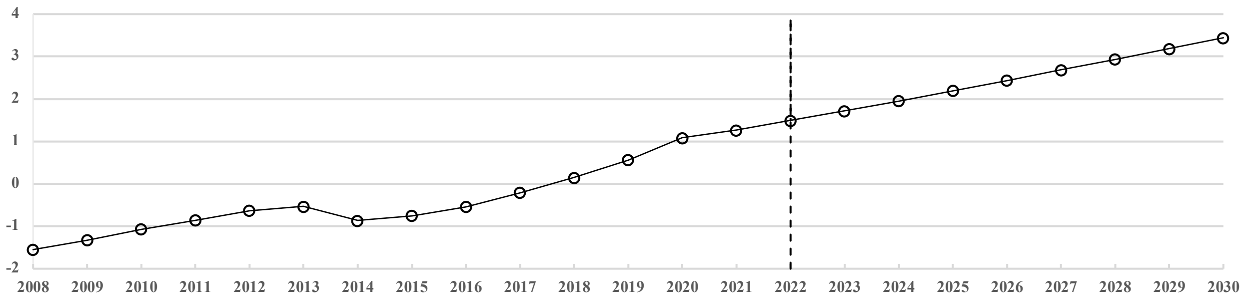
C. Dynamic Factor Model estimation and results

C1. Estimation Results

C1.1. Sector – input

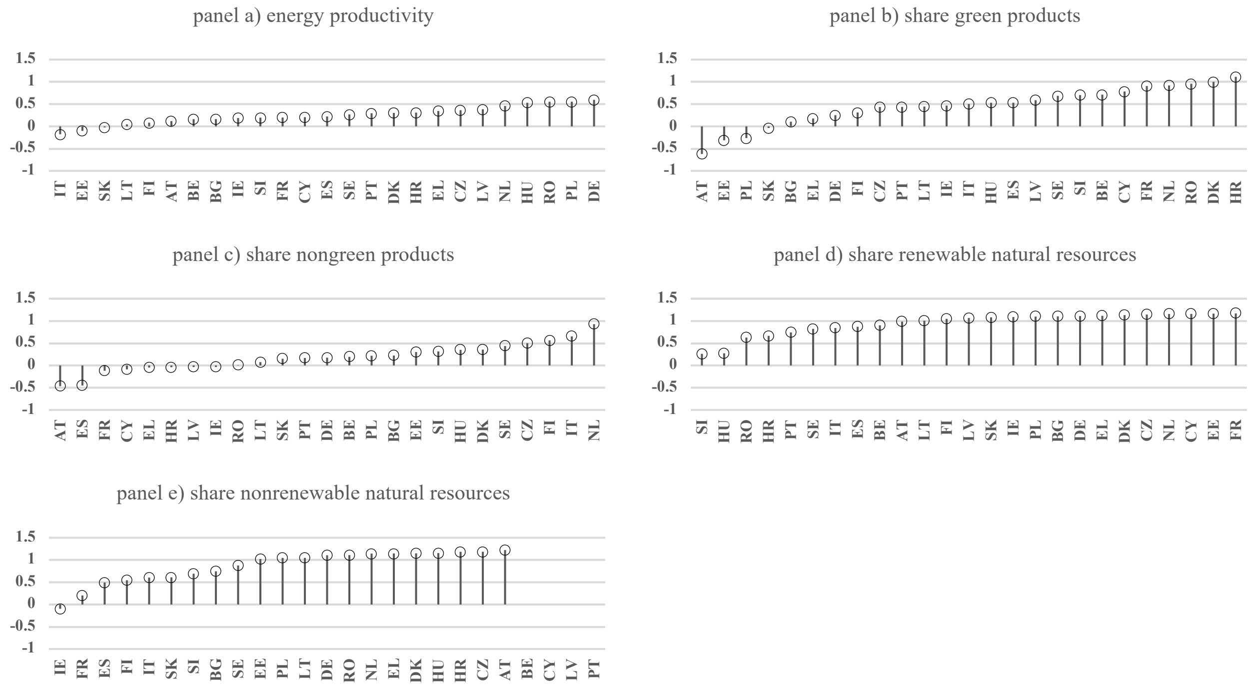
Energy

Figure 50: Green factor – Energy



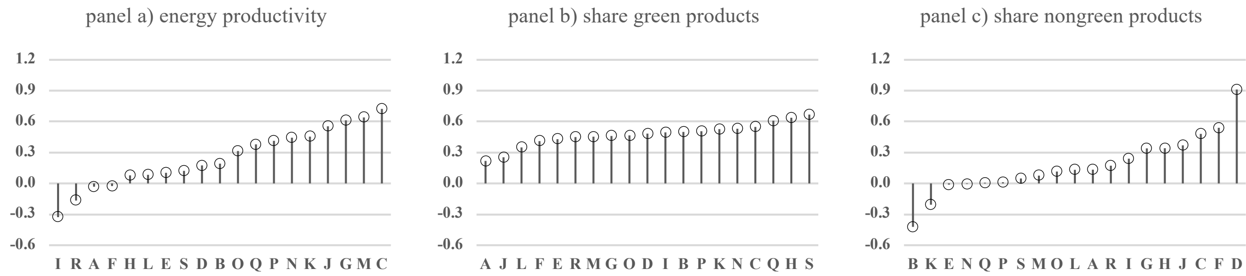
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 51: Average loadings – Energy – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

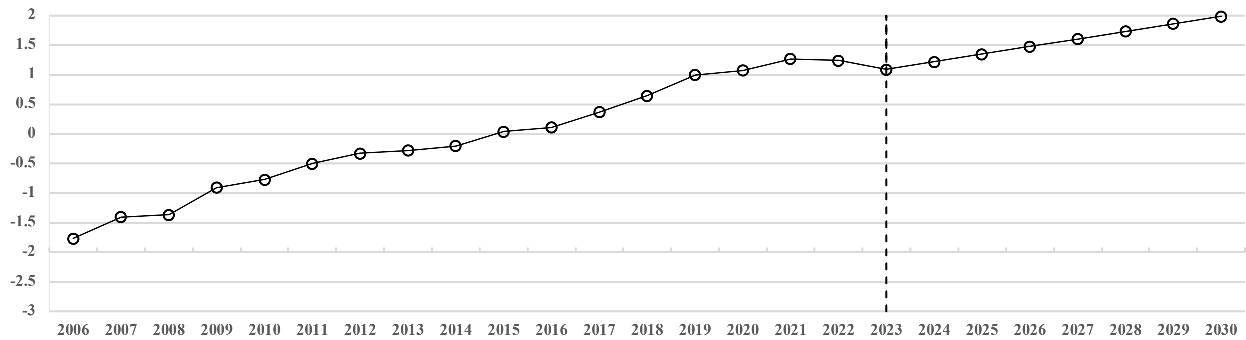
Figure 52: Average loadings – Energy – Sectors



Source: Own elaboration.
Note: Loadings averaged over MS. See https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

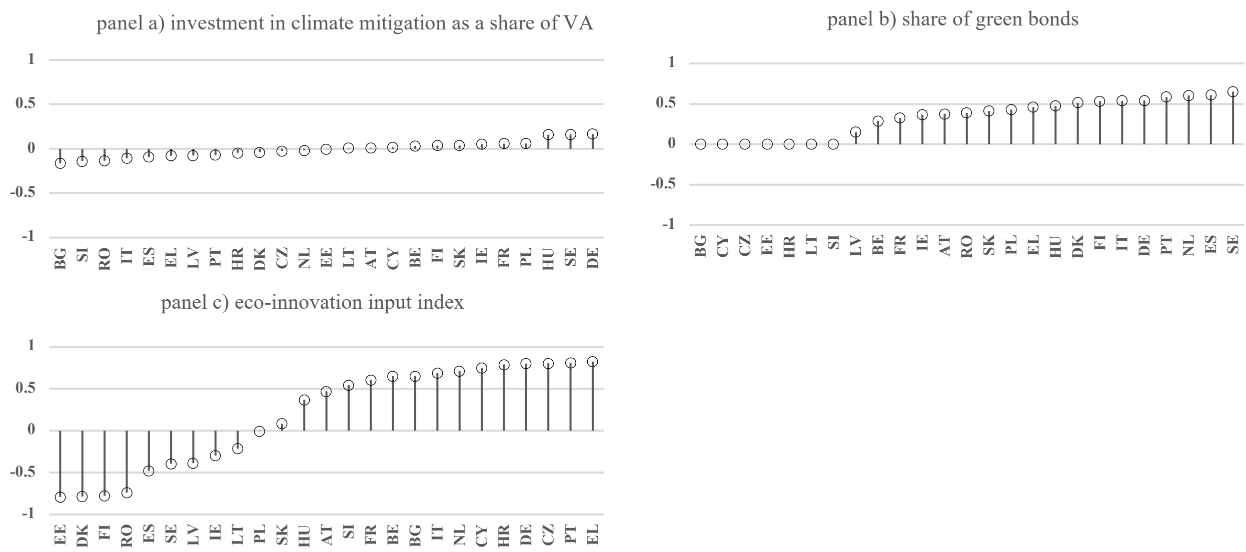
Capital

Figure 53: Green factor – Capital



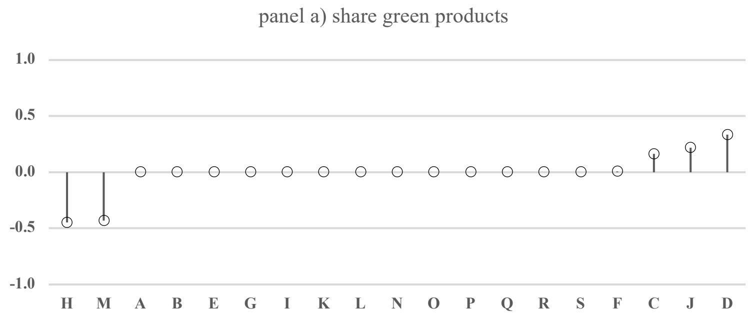
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 54: Average loadings – Capital – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

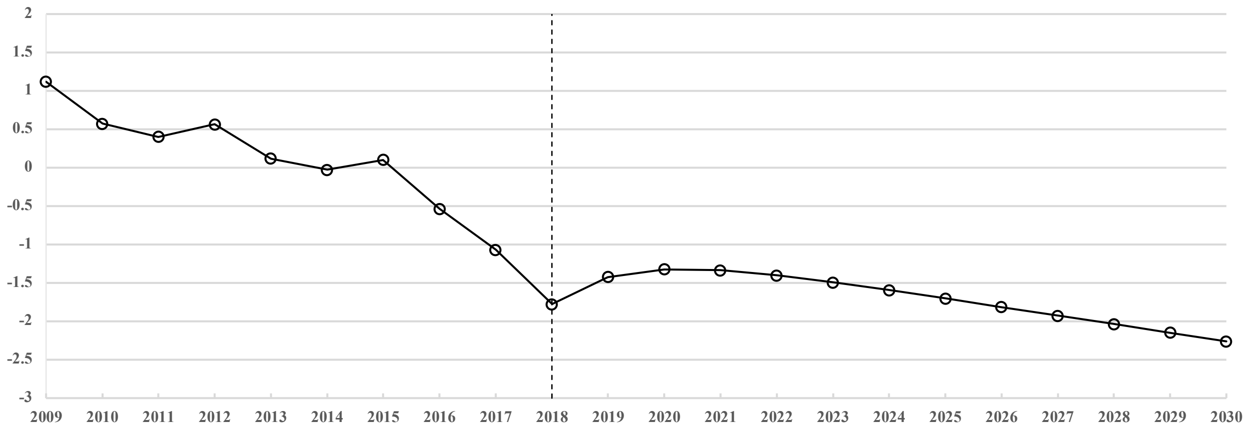
Figure 55: Average loadings – Capital – Sectors



Source: Own elaboration.
Note: Loadings averaged over MS. See https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

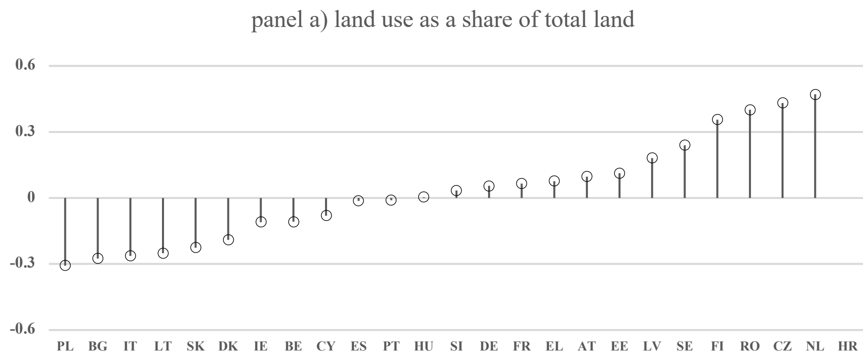
Land use

Figure 56: Green factor – Land use



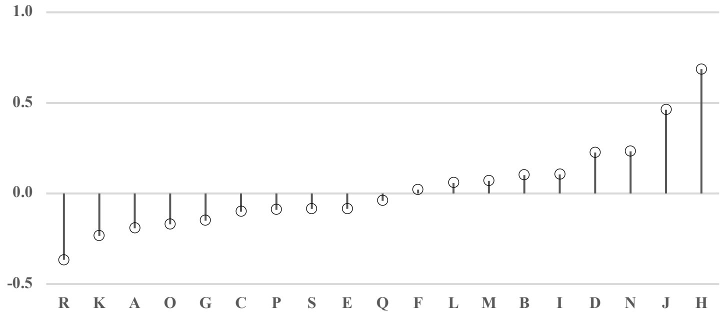
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 57: Average loadings – Land use – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

Figure 58: Average loadings – Land use – Sectors

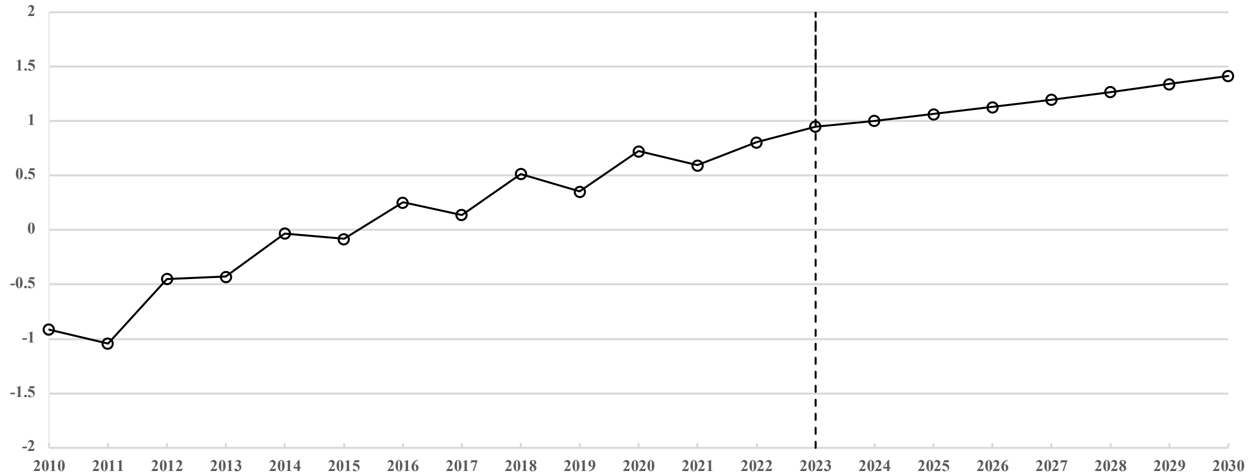


Source: Own elaboration.
Note: Loadings averaged over MS.
https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

C1.2. Sector – process

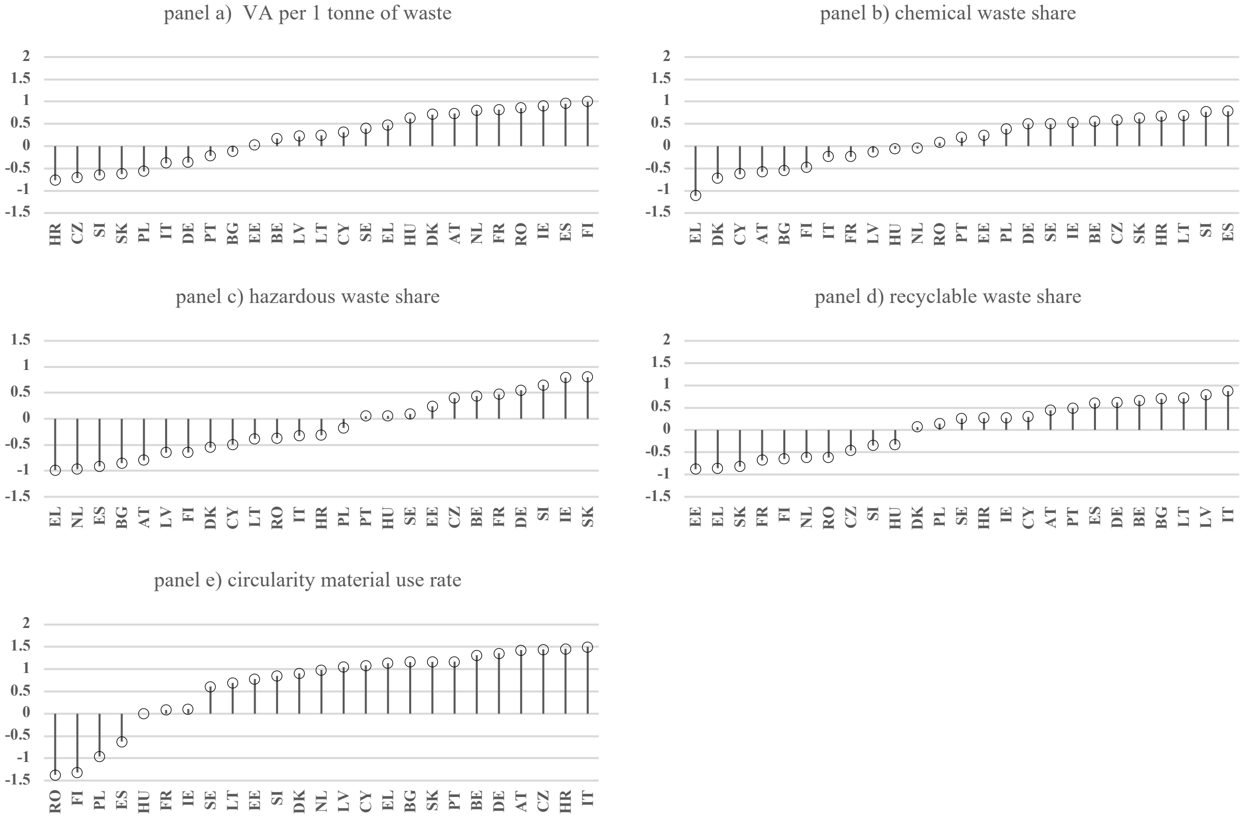
Waste and circularity

Figure 59: Green factor – Waste and circularity



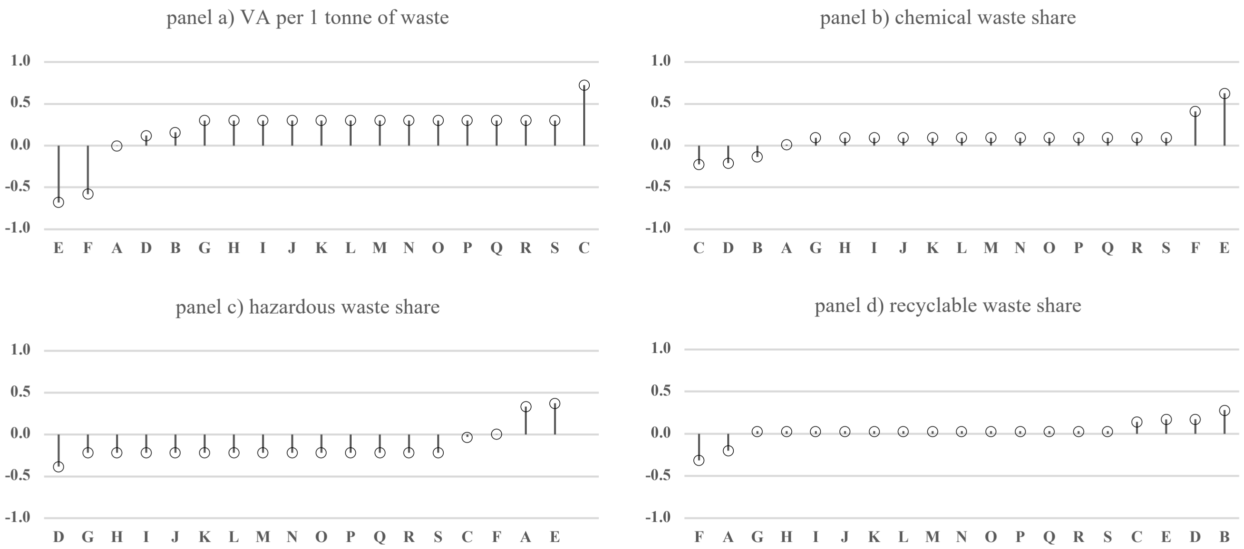
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 60: Average loadings – Waste and circularity – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

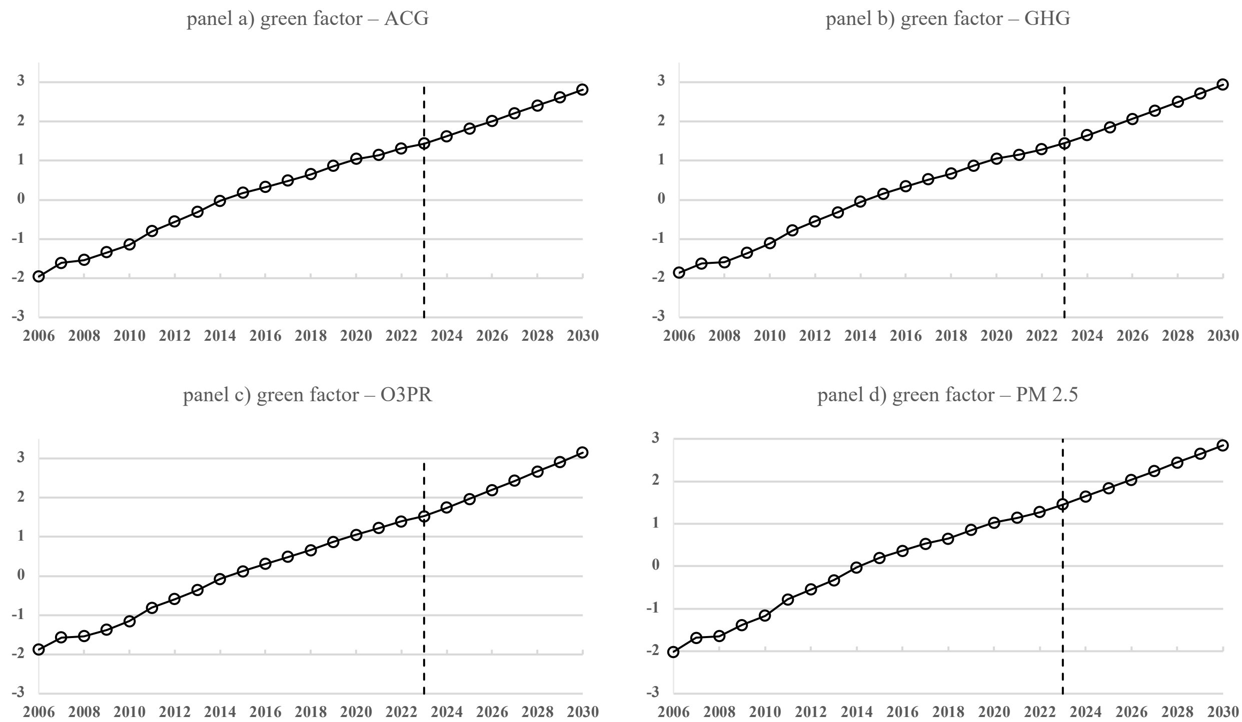
Figure 61: Average loadings – Waste and circularity – Sectors



Source: Own elaboration.
Note: Loadings averaged over MS. See https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

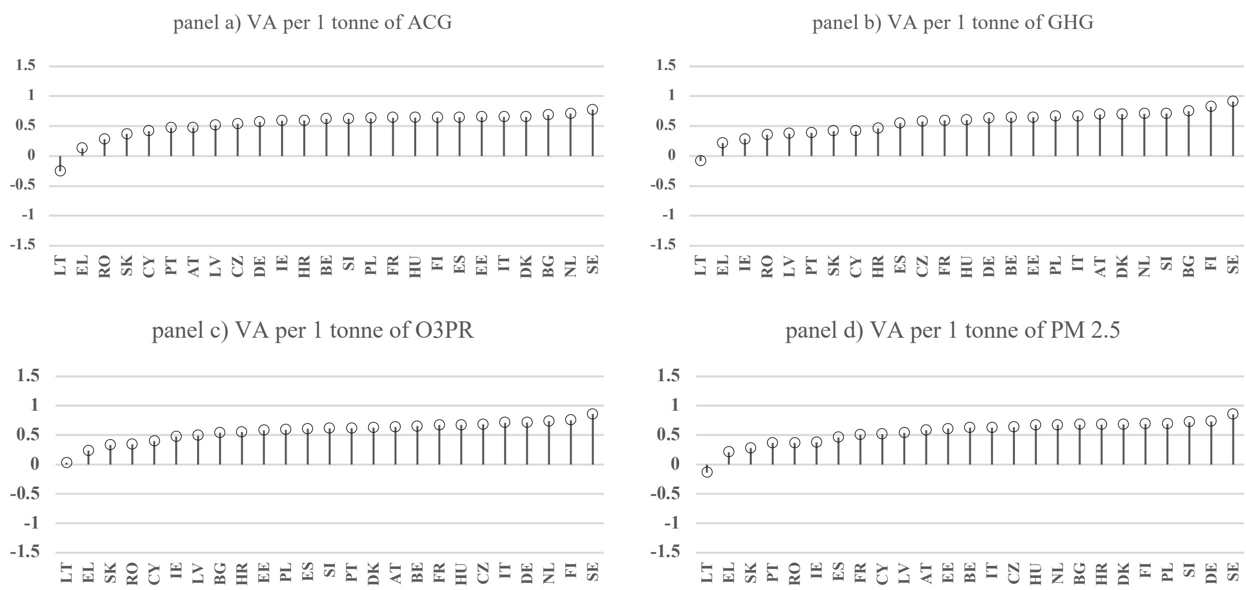
Air emissions

Figure 62: Green factor – Air emissions



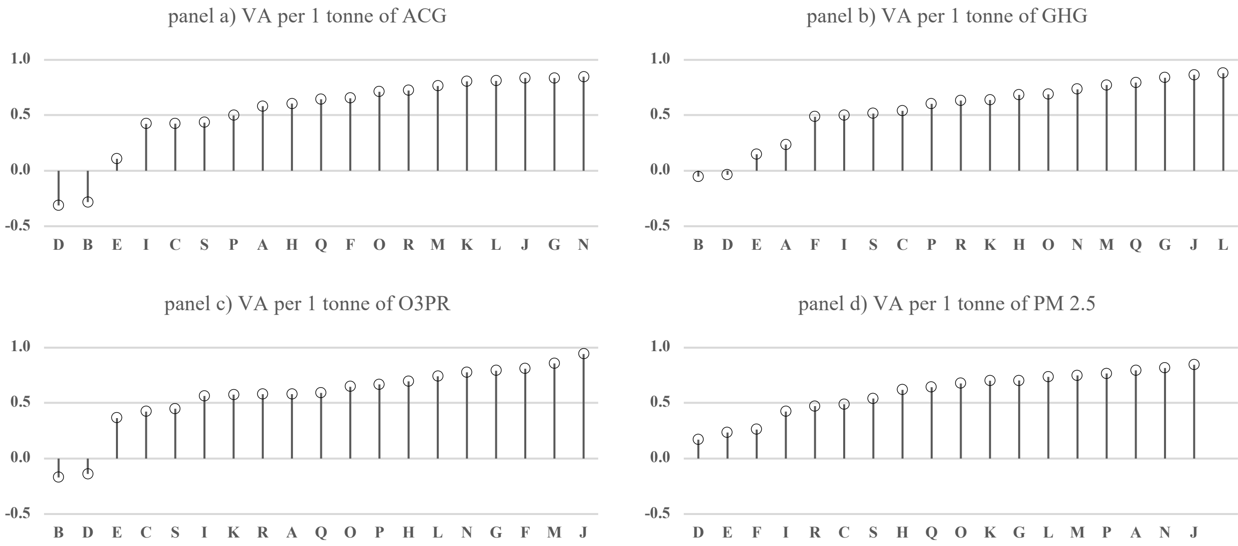
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 63: Average loadings – Air emissions – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

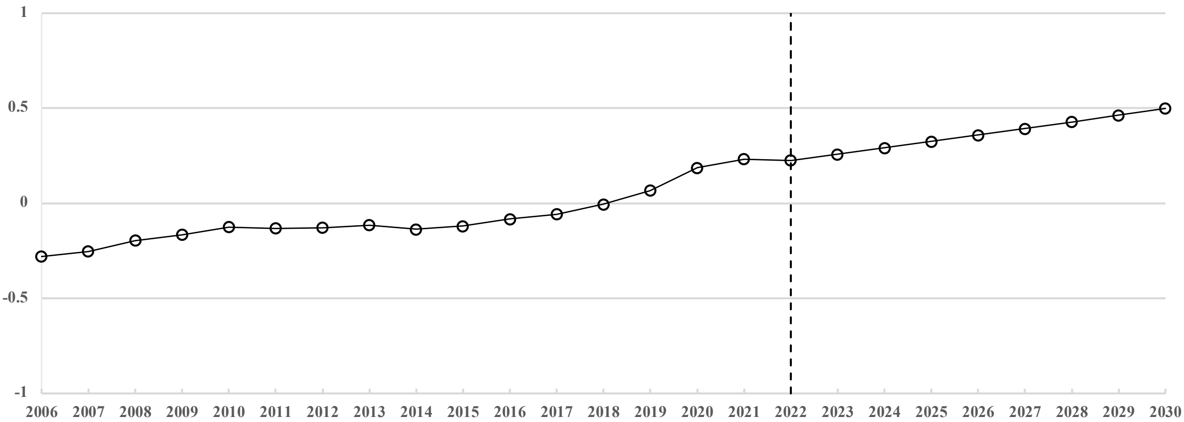
Figure 64: Average loadings – Air emissions – Sectors



Source: Own elaboration.
Note: Loadings averaged over MS. See https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

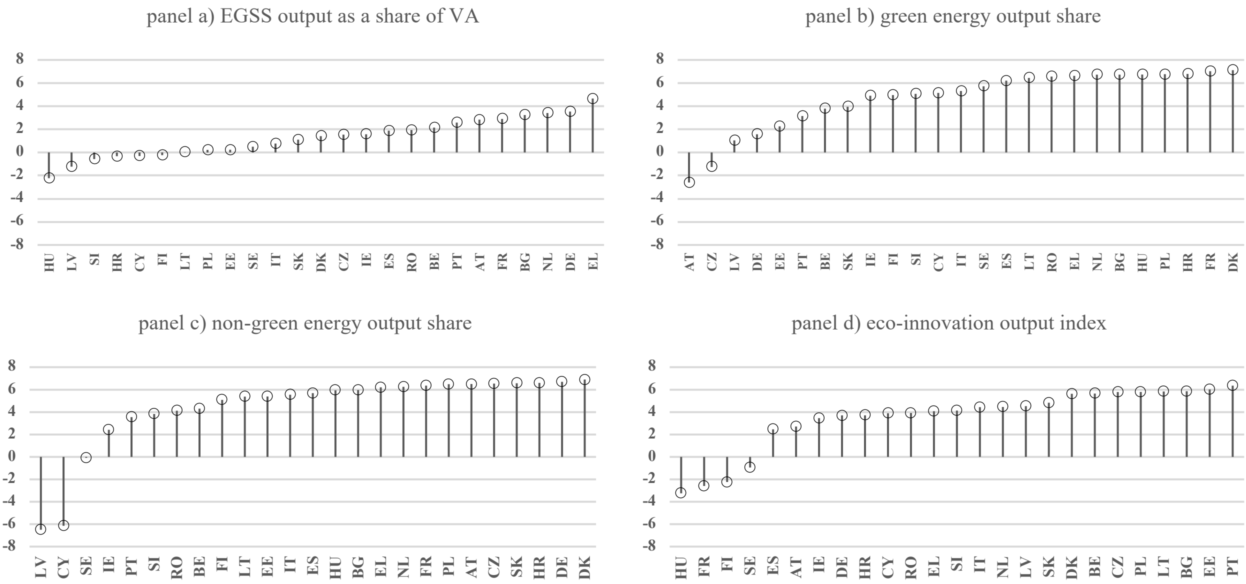
C1.3. Sector – output

Figure 65: Green factor – Output



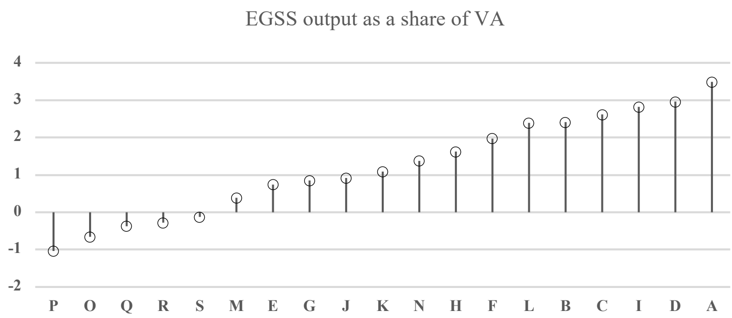
Source: Own elaboration.
Note: Factor extracted using the Kalman smoother. The dashed vertical line indicates the year after which the factor is forecasted.

Figure 66: Average loadings – Output – Member States



Source: Own elaboration.
Note: Loadings averaged over sectors.

Figure 67: Average loadings – Output – Sectors



Source: Own elaboration.
Note: Loadings averaged over MS. See https://en.wikipedia.org/wiki/NACE_rev2 for the list of sector codes.

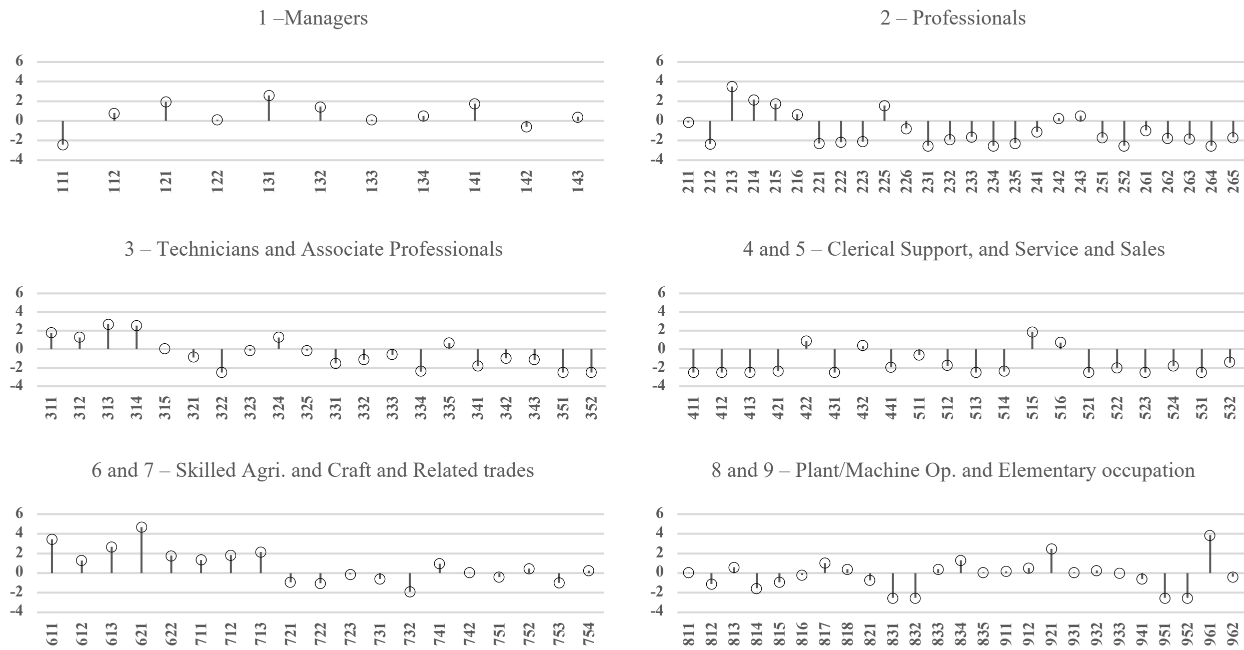
D. PCA ON SKILLS INDICATORS

Table 5: PCA results on skills and knowledges indicators

	Descriptive statistics		PCA			KMO	
	mean	sd	Loading	Eigenvalue	Share of the variance	Variable	Overall
Total green knowledges and skills	1.46	0.83	0.60			0.61	
Essential Green Knowledges	1.70	1.24	0.56	2.56	0.89	0.75	0.68
Essential Green skills	1.08	0.89	0.57			0.70	

Source: Own elaboration.

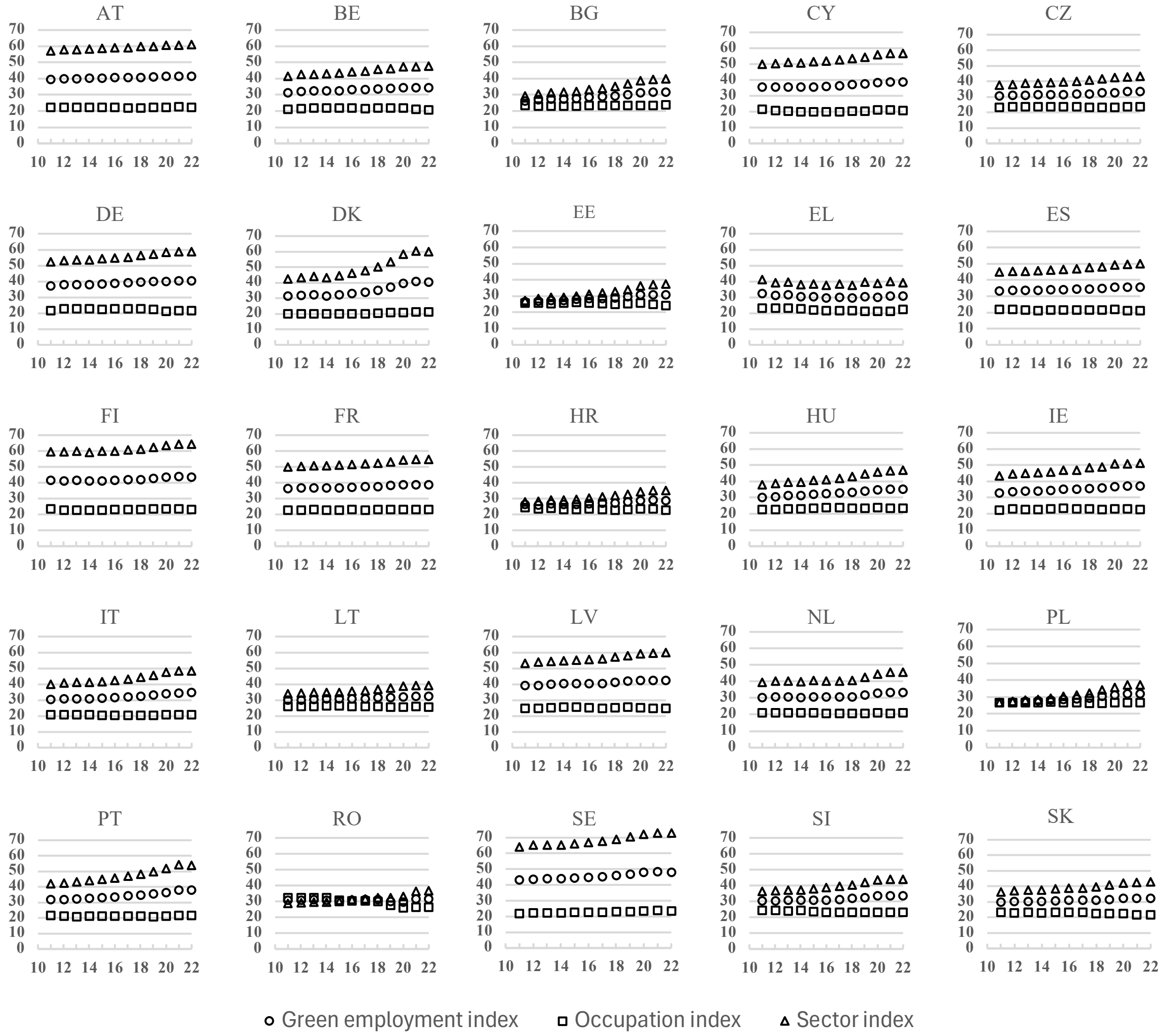
Figure 68: PCA scores on skills and knowledges indicators



Source: Own elaboration.
Note: PCA scores. https://ilostat.ilo.org/classification-occupation/

E. Green Employment Index

Figure 69: GEI, sector and occupation indices.



Source: Own elaboration.