Do common shocks drive changes in aggregate emissions intensity?*

Xiyu Ren[a,b*], Fulvia Marotta[a,c], and François Lafond[a,b]

[a]Smith School of Enterprise and the Environment, University of Oxford
[b]Institute for New Economic Thinking at the Oxford Martin School, University of Oxford
[c]De Nederlandsche Bank

Abstract

In the UK, aggregate emissions intensity has declined by about a factor of two over the last three decades. Prior research attributes most of this decline to reductions within industries rather than shifts in the composition of economic activity. This paper investigates whether such within-industry progress primarily reflects industry-specific factors or common forces operating across industries. Using a newly constructed panel of UK industry-level GHG emissions and gross value added for 1990–2022, we estimate a block-level dynamic factor model that decomposes changes in emissions intensity into global, block-level, and idiosyncratic components. We find that industry-specific factors account for the majority of variation in emissions intensity changes, though common shocks, either global or at the level of groups of industries, play a smaller but non-negligible role. We further show how patterns of co-movement partly reflect the way that emissions are recorded at the activity level and allocated to industries, a feature with implications for interpreting industry-level decarbonization dynamics.

Keywords: Emissions intensity; Sectoral heterogeneity; Dynamic factor models; Climate policy; Environmental macroeconomics; Directed technical change
JEL Codes: L16, C38, E32, C32, Q54, O33

*This is the Author Accepted Manuscript of `https://doi.org/10.1016/j.jeem.2026.103384`. We would like to thank Matthew Agarwala, Vasco Carvalho, Jennifer Castle, Angelo Cuzzola, Sam Fankhauser, Bas Heerma van Voss, Rasmus Lönn, Leonardo Melosi, Émilien Ravigné, Michele Piffer and the audiences of seminars at Oxford (Climate Econometrics), Utrecht University, European Central Bank, De Nederlandsche Bank, Bank of England, LSE, Scuola Superiore Sant'Anna, the CFE conference 2024, the Econometric Modelling of Climate Change conference 2024, IAAE conference 2025. We gratefully acknowledge funding from UK Research and Innovation through Economic and Social Research Council grant PRINZ (ES/W010356/1), Oriel college, Baillie Gifford, and from the Institute for New Economic Thinking at the Oxford Martin School. Corresponding author: f.marotta@dnb.nl

1 Introduction

Understanding the drivers of emissions reductions is essential for designing effective climate policy. While much of the existing literature focuses on aggregate, country-level emissions trends, a growing body of research has begun to explore industry-level dynamics, recognizing that production processes – and thus mitigation opportunities – are inherently sector-specific (e.g. Agarwala and Martin, 2022; Rafaty et al., 2025). However, most of this work examines only a limited set of sectors or case studies, and systematic, economy-wide analyses of industry-level emissions remain rare. At the same time, climate policy frameworks are becoming increasingly granular. Net-zero commitments and transition plans are now formulated not only at the national level, but also by sector and even by individual firms (Rogelj et al., 2021; Fankhauser et al., 2022; GFANZ, 2022). This growing emphasis on sectoral targets raises a fundamental question: when we observe emissions reductions, to what extent do they reflect sector-specific choices and circumstances, and to what extent are they driven by broader, economy-wide forces?

This distinction matters because firm- and industry-level emissions targets are most meaningful when the entities themselves can directly influence decarbonization outcomes, for instance by investing in cleaner technologies or adjusting their input mix. From this bottom-up perspective, aggregate emissions reductions are viewed as the sum of industry-specific actions. However, industry-level changes may also reflect broader, economy-wide forces rather than deliberate decisions. For example, an industry might reduce its emissions by actively downsizing its vehicle fleet, or it could simply benefit from an economy-wide shift in the vehicle stock from petrol to electric, regardless of its own choices. In the first case, the emissions drop reflects a conscious decision; in the second, the industry is a passive beneficiary of systemic change.

This distinction is not merely academic. Emissions reductions observed at the industry level often reflect broader systemic forces rather than deliberate actions by firms themselves. For example, Germany's industrial emissions fell in 2024 largely due to a mild winter and an economic slowdown, rather than targeted decarbonization within firms or sectors.[1] Similarly, during the COVID-19 pandemic, global CO_2 emissions dropped by as much as 17% in early April 2020 relative to 2019 levels, almost entirely due to reductions in industrial and transport activity caused by lockdowns (Le Quéré et al., 2020). These examples illustrate why distinguishing passive, system-driven reductions from proactive, industry-led decarbonization is critical for both measurement and policy.

In this paper, we ask: to what extent are reductions in aggregate emissions intensity driven by sector-specific changes versus common, economy-wide forces? We take a bottom-up perspective, analyzing emissions intensity trends across industries while paying close attention to patterns of co-movement. Although many studies focus on aggregate drivers of decarbonization, there is limited quantitative evidence on how industry-level dynamics are driven by economy-wide mitigation efforts, and how these dynamics shape the aggregate outcome. A notable study is Rafaty et al. (2025), who introduce interactive fixed effects (effectively, a factor model) as controls in their study of the impact of carbon prices on emissions, but do not provide additional discussions of the factors

[1] See, for example, Reuters (2025): "German emissions fell 3% in 2024 due to economic weakness, think tank says."

and loadings. We address this gap using a newly constructed panel dataset of UK greenhouse gas emissions and output (value added), covering 106 industries over the period 1990 to 2022. This rich dataset allows us to track emissions intensity across sectors and over time, enabling a systematic decomposition of sectoral versus aggregate trends.

We begin by documenting three stylized facts that motivate our empirical strategy. First, the decline in aggregate emissions intensity is primarily driven by reductions within industries, rather than by a reallocation of economic activity away from more polluting sectors. This result, consistent with Agarwala and Martin (2022), supports a bottom-up perspective: understanding aggregate trends largely requires analyzing industry-level changes. Much of the existing literature in energy economics has built on similar decompositions, often extending them into increasingly detailed components (see, for example, Wen et al. (2022); Hardt et al. (2018), and the early survey by Ang and Zhang (2000)). Our approach diverges from this tradition. Rather than further disaggregating the aggregate trend into additional factors, we focus on the "within-industry" component itself and examine whether its dynamics are driven primarily by idiosyncratic industry shocks or by broader common forces. Second, although emissions are concentrated – 10 industries account for just under three-quarters of total emissions – nearly all industries contribute some emissions. This underscores the importance of analyzing emissions patterns across the full industrial distribution. Third, we find strong evidence of co-movement in emissions intensity—both economy-wide and within subsets of industries. In particular, we observe a striking cluster of synchronized trends among service industries. This co-movement suggests the presence of common or block-level drivers, which motivates our use of a block-level dynamic factor model.

A simple approach to separate common and idiosyncratic shocks is to compute cross-sectional averages, interpreting the average as a common shock and the residual as an idiosyncratic shock (see e.g., Gabaix 2011). Factor models provide a richer structure, by allowing for heterogeneous loadings, that is, enabling different industries to respond differently to the common shocks (Barigozzi et al., 2024). However, simple factor models assume that while every unit may respond differently to the latent factor, this latent factor is global.

Our approach extends standard factor models by introducing block-level factors, in addition to a global economy-wide common factor. This allows us to distinguish between shocks that affect all industries to varying degrees and those that affect subsets of industries (e.g., services or manufacturing) with shared characteristics. The decomposition into global, block, and idiosyncratic components follows the framework of Delle Chiaie et al. (2022), and has been used widely in empirical studies of economic and financial co-movements (e.g., Bańbura et al. 2011, Forni and Reichlin 1998, Kose et al. 2003, Miranda-Agrippino and Rey 2020, Mumtaz and Theodoridis 2017, Mumtaz and Musso 2021). To our knowledge, this is the first study to apply a block-level dynamic factor model to emissions intensity. This extension is not just methodological: relying on a standard factor model can mislead inference by attributing localized co-movement – where only certain groups of industries move together – to a single global factor. Such misattribution risks overstating the role of economy-wide forces and understating the importance of sector-specific dynamics.

This matters for interpretation. A global common factor in our model may reflect a genuinely aggregate shock in the economy – such as a widely applicable carbon price, a major shift in a universal input (such as transportation systems), or a broad energy-saving campaign. These shocks are expected to influence all industries, albeit to varying degrees (as captured by heterogeneous loadings). In contrast, block-level common shocks may arise from policies or shocks that affect only a segment of the economy, such as targeted subsidies for manufacturing or regulation of specific service activities. By introducing block-level factors, our model captures this intermediate structure of co-movement, enabling a more accurate and interpretable decomposition of emissions intensity dynamics. Although we estimate our factor model in first differences to avoid any concern about spurious factors (Onatski and Wang, 2021), so that as a result we do not capture common trends (Kaufmann and Strachan, 2024), the extent to which the common factor drives the overall dynamics still provides us with an estimate of the importance of macro shocks and universal inputs.

To understand the origins of co-movement in emissions intensity, it is useful to consider how emissions are measured and attributed to industries. Emissions are first recorded at the level of specific economic activities, such as combustion, chemical processing, and transport. These activity-level emissions are then allocated to industries based on their reliance on those activities. Consequently, the co-movement observed across industries can arise for two reasons: first, because the emissions intensity of certain activities co-moves over time, and second, because multiple industries share the same activities as upstream inputs and services.

This bipartite activity-industry structure creates a natural basis for both global and block-level common components. Some activities, such as road transport, are pervasive and contribute to emissions in all industries, generating global co-movement. Others, like chemical processing or metal refining, are linked to narrower sets of industries and give rise to block-level dynamics. In our context, emissions shocks associated with shared activities can induce synchronized changes in emissions intensity across downstream sectors, even if their final output differs. This structural mapping helps explain and validate the layered decomposition captured by our factor model.

We find that idiosyncratic, industry-specific shocks explain the majority of variation in emissions intensity. This suggests that decarbonization has largely been driven by distinct, industry-specific responses. Yet, we identify that both a global common factor and block-level factors explain a non-negligible share of the variance. The global factor features loadings much more pronounced in services industries, and likely reflects broad and aggregate forces, such as clear policy signals towards a decarbonization agenda, but also changes in the carbon intensity of "universal" inputs, used by all industries, such as road transport. The block-level factors can explain a substantial amount of variance, particularly in the Agriculture, Energy, and Chemicals blocks.

Our findings connect to several strands of the literature. First, they speak to research on the role of micro-heterogeneity, idiosyncratic shocks, and common factors in shaping aggregate outcomes (e.g., Gabaix 2011; Foerster et al. 2011), by showing that most variation in emissions intensity reflects industry-specific shocks rather than pervasive economy-wide forces. Second, they relate to work on the cyclical dynamics of emissions (e.g., Doda 2014; Khan et al. 2019), as we document that short-run fluctuations in emissions intensity are driven more by industry-level disturbances than by common macroeconomic shocks. Third, our decomposition complements studies of emissions in networked production systems (e.g., Van Den Ende et al. 2025), as the block-level factors we identify can be viewed as capturing co-movement arising from shared inputs and upstream activities.

Finally, our results inform debates on the design of net-zero policy frameworks (e.g., Stechemesser et al. 2024; Colmer et al. 2024; Kaartinen and Prane 2024), by clarifying the extent to which sectoral emissions reductions reflect targeted industry action versus broader systemic forces.

The patterns we uncover, ranging from strong industry-specific dynamics to block-level comovement and a growing common component, pose real challenges for empirical modeling. Emissions arise from a rich and partly unobserved network of shared activities, and industries exhibit heterogeneous responses to shared trends. Our goal is to strike a balance between capturing this empirical complexity and maintaining a tractable, interpretable framework. We return to these modeling trade-offs and their limitations in the discussion.

The paper is structured as follows. Section 2 introduces the novel dataset on emissions intensities and presents key stylized facts that motivate our empirical strategy. Section 3 outlines the methodology, including the allocation-based data-generating process and the block-level dynamic factor model. Section 4 presents the main findings, including the decomposition of industry-level variation, evidence of synchronization in emissions reduction, and the contribution of each component to aggregate emissions intensity dynamics. Section 6 concludes and discusses implications for climate policy and future research.

2 Data and stylized facts

2.1 Data sources

We employ industry-level data on greenhouse gas (GHG) emissions and gross value added (GVA) for the United Kingdom over the period 1990–2022. Emissions data are sourced from the Office for National Statistics (ONS, 2022), while GVA data are obtained from the ONS GDP Output Approach Low-Level Aggregates dataset (ONS, 2023a). The datasets are harmonized at the most disaggregated level available, yielding 106 industries classified into 19 one-digit Standard Industrial Classification (SIC) categories[2]. Each industry approximately corresponds to a two-digit SIC07 category (see Supplementary Information, Appendix A.2.1).

We choose GVA rather than Gross Output. Emissions are more likely to be a function of gross output than value added, so one could use gross output. However, GVA is more readily available at the right level of granularity for the entire period, and, more importantly, the gross output of various industries is mechanically correlated due to input-output relationships, i.e., a given unit of production counts as gross output of several industries along the supply chain. A disadvantage of using GVA is that it can approach zero (as we will see below for coal mining, comparing the two panels of Fig. 2).

GHG emissions are measured on a "scope one" basis, representing direct emissions from sources

[2] While using a different aggregation level or aggregating specific industries together could be justified, here we take the simplest and most transparent approach that exploits as much cross-sectional information as possible to speak to the question of common factors in emissions intensity data.

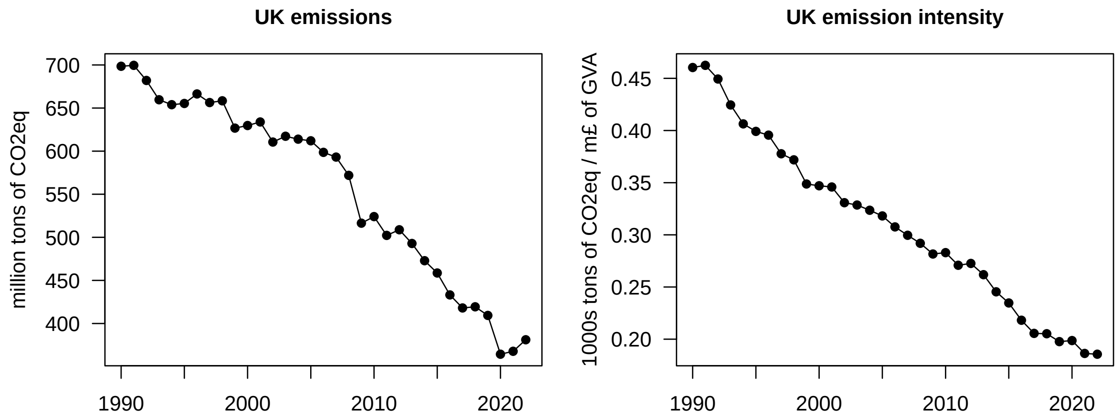

Figure 1: Aggregate emissions and emissions intensity

owned or controlled by the industry. We use scope 1 emissions because this avoids double counting of emissions, and because this is much simpler. Of course, this also implies our results do not speak to the offshoring of emissions that is documented in the literature on production vs consumption-based emissions (see e.g. Hardt et al. (2018)).

Emissions are calculated on both a production and residence basis, capturing emissions generated by UK residents and UK-registered businesses, irrespective of geographic location. This "enables direct comparison of emissions by sector of UK industry and households with main economic indicators including gross value added" (ONS, 2025).

Emissions are reported in kilotonnes of CO_2 equivalent ($ktCO_2eq$) and GVA is expressed in chained volume measures (CVMs). We define emissions intensity as the ratio of GHG emissions to GVA, expressed in $ktCO_2eq$ per million pounds sterling.

The dataset captures approximately 75–80% of total greenhouse gas (GHG) emissions and 90% of gross value added (GVA) in the United Kingdom. Both datasets exclude industry "T" (households) and other consumer-related emissions (divisions 100 and 101), which collectively account for roughly 20–25% of total emissions but less than 1% of total GVA. Industries for which GHG or GVA observations are zero or unavailable are also excluded. These include division 07 (Mining of Metal Ores) and division 65.3 (Pension Funding), which both record zero GVA,[3] as well as division 68.2 (Imputed Rental), which generates zero emissions but represents roughly 10% of total GVA.

Figure 1 presents the aggregate evolution of UK GHG emissions and emissions intensity over 1990–2022. Aggregate emissions declined by roughly a half over this period, while real GDP expanded, resulting in a more than 50% reduction in emissions intensity. Notably, emissions display considerably greater cyclical variation around their long-term trend than emissions intensity. This reflects the close association between emissions and economic activity: for instance, emissions contracted sharply in 2020 during the COVID-19-induced recession, whereas emissions intensity

[3] Division 65.3 (Pension Funding) accounts for approximately 0.001% of total emissions, while division 07 accounts for about 0.1%.

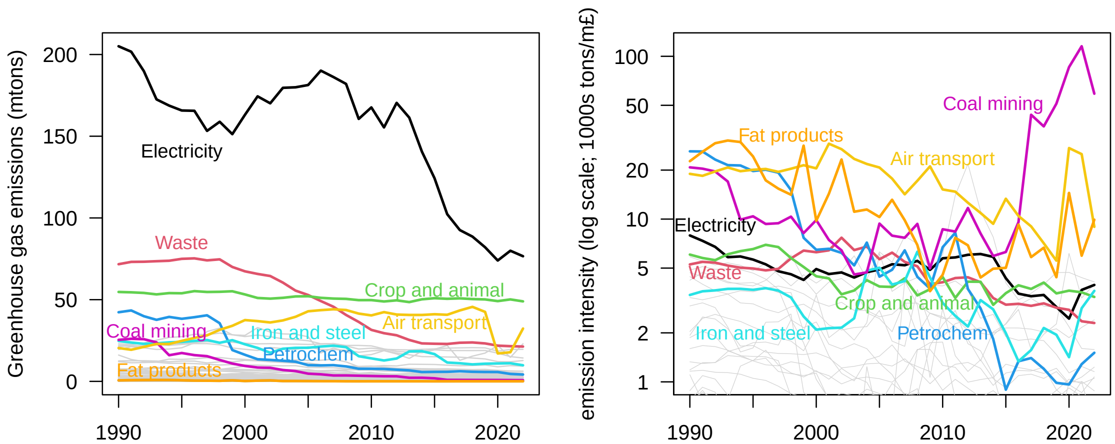

Figure 2: Industrial GHG and GHG intensity, with selected industries shown in colours, with manually shortened labels. Two values for emissions intensity are negative (GVA can be negative) and replaced by 1 for plotting purposes, and the y-axis range clips values below 1.

continued to decline broadly in line with its pre-existing trend.

2.2 *Fact 1*: Emissions are concentrated in a few industries

To better understand the sources of the aggregate decline in emissions and emissions intensity, Figure 2 plots total emissions and emissions intensity by industry. The left panel illustrates that a relatively small number of industries account for the majority of total emissions, while the remaining industries contribute comparatively little to aggregate emissions.

To evaluate the extent of concentration more precisely, we calculate the proportion of GHG and GVA contributed by each industry in relation to the total emissions and value added of the economy. Table 1 shows the cumulative percentage of GHG and GVA of the top 10 emitters, showing that they collectively account for more than 70% of the total emissions in the economy, but only 12% of GVA. This concentration suggests that idiosyncratic shocks to a few key industries could account for a large share of aggregate emissions dynamics—a pattern consistent with the idea of a granular origin of aggregate fluctuations, where shocks to large or dominant sectors can shape macro-level outcomes (Gabaix, 2011). One particularly salient example is the electricity generation industry, which experienced a pronounced decline in emissions over the sample period. As has been well documented (Hendry, 2020), this trend largely reflects the phasing out of coal in power generation – a clear case of sector-specific technological change.

The waste and petrochemical industries display a pattern similar to electricity generation, with strong declines in emissions largely driven by reductions in emissions intensity. In contrast, sectors such as crop and animal production, fat product manufacturing, iron and steel, and air transport exhibit no consistent trend in either emissions or emissions intensity. The impact of COVID-19 is also visible in the data: emissions from air transport dropped sharply in 2020 due to reduced

Average from 1990 to 2022	% GHG	%GVA	GHG/GVA
Electricity generation	26.98	1.68	4.98
Crop and animal	9.58	0.65	4.49
Waste	8.24	0.56	4.72
Air transport	6.50	0.13	17.62
Extraction of petroleum and gas	4.10	6.77	0.26
Land transport	3.98	1.38	0.87
Water transport	3.94	0.38	3.32
Manufacture of petroleum products	3.34	0.20	6.03
Basic iron and steel	3.28	0.35	3.13
Petrochemicals	2.82	0.14	8.74
Sum top 10 emitters	72.76	12.24	

Table 1: Proportion of emissions and GVA of the industries out of the economy and their emissions intensity

flight activity, but emissions intensity increased, reflecting lower passenger or cargo loads and, consequently, less value added per flight.

The concentration of emissions in a few sectors makes it plausible that aggregate trends may be shaped by a handful of dominant industries. However, even if this is the case, these key industries may be driven by common factors. We will return to this question throughout the paper, but before, we need to formally evaluate the contribution of composition effects to the trend in aggregate emissions intensity.

2.3 ***Fact 2*: Industry-specific emission intensity improvements have been more important than reallocation effects**

The decline in aggregate emissions intensity can come from two sources: individual industries are becoming cleaner (lower intensity), or dirty industries are becoming relatively smaller (keeping the same intensity, but representing a lower share of value added). To see this, let us write aggregate emissions intensity as the sum of industry-level emissions,

$$\mathcal{E}_t \equiv \frac{E_t}{Y_t} = \frac{\sum_i E_{it}}{Y_t} = \sum_i w_{it}\mathcal{E}_{it}, \tag{1}$$

where E_t and Y_t are the aggregate emissions and GVA at time t, E_{it} and Y_{it} are the industry-level emissions and GVA at time t, and $w_{it} \equiv Y_{it}/Y_t$ are the shares of industry i in total GVA. Now, using Eq. 1, we can decompose the *change* in emissions intensity into a “within” and a “between” effect as

$$\Delta\mathcal{E}_t \equiv \mathcal{E}_t - \mathcal{E}_{t-1} = \underbrace{\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it}}_{\text{“within effect”}} + \underbrace{\sum_i \bar{\mathcal{E}}_{it}\Delta w_{it}}_{\text{“between effect”}}, \tag{2}$$

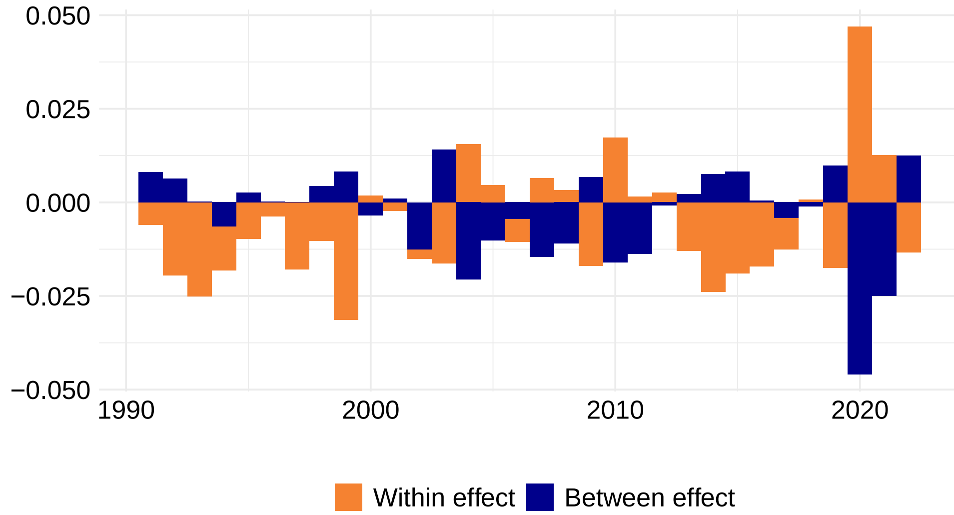

Figure 3: Within and between effect. The figure demonstrates that in most years, the majority of the aggregate emissions intensity change is attributed to the within effect.

where we have used the weights $\bar{w}_{it} \equiv (w_{i,t} + w_{i,t-1})/2$ and $\bar{\mathcal{E}}_{it} \equiv (\mathcal{E}_{i,t} + \mathcal{E}_{i,t-1})/2$ to avoid the presence of a covariance-like term, see Griliches and Regev (1992).

The within effect reflects changes in emissions intensity at the industry level, holding industry value-added shares constant, while the between effect captures the contribution of shifts in GVA shares across industries with differing intensities. The relative importance of these two components depends on the level of industrial aggregation in the data. For example, in our classification, electricity generation from coal and gas is grouped into a single industry, so substitution between coal and gas is recorded as part of the within effect rather than the between effect.

Figure 3 plots the year-by-year contributions of the within and between effects to changes in aggregate emissions intensity. The within effect drives most of the decline over 1990–2022, remaining negative in nearly all years. The between effect is smaller and fluctuates, occasionally offsetting or amplifying the within effect. In 2020, the between effect dominates, as the COVID-19 shock caused a sharp contraction in contact- and emission-intensive industries, temporarily accelerating the aggregate decline in emissions intensity.

The dominance of the within effect in driving declines in aggregate emissions intensity is consistent with prior findings. Amadei et al. (2024), for instance, document that within-sector energy efficiency improvements account for most reductions in energy intensity across OECD countries, with structural shifts across sectors playing a smaller role. A large literature in energy economics builds increasingly detailed decompositions – typically using energy rather than emissions data – to disentangle such effects. These studies identify multiple mechanisms, such as reductions in useful energy per unit of output, improved conversion of final to useful energy, or shifts in domestic consumption versus offshoring of production. For the UK, Hardt et al. (2018) find that both

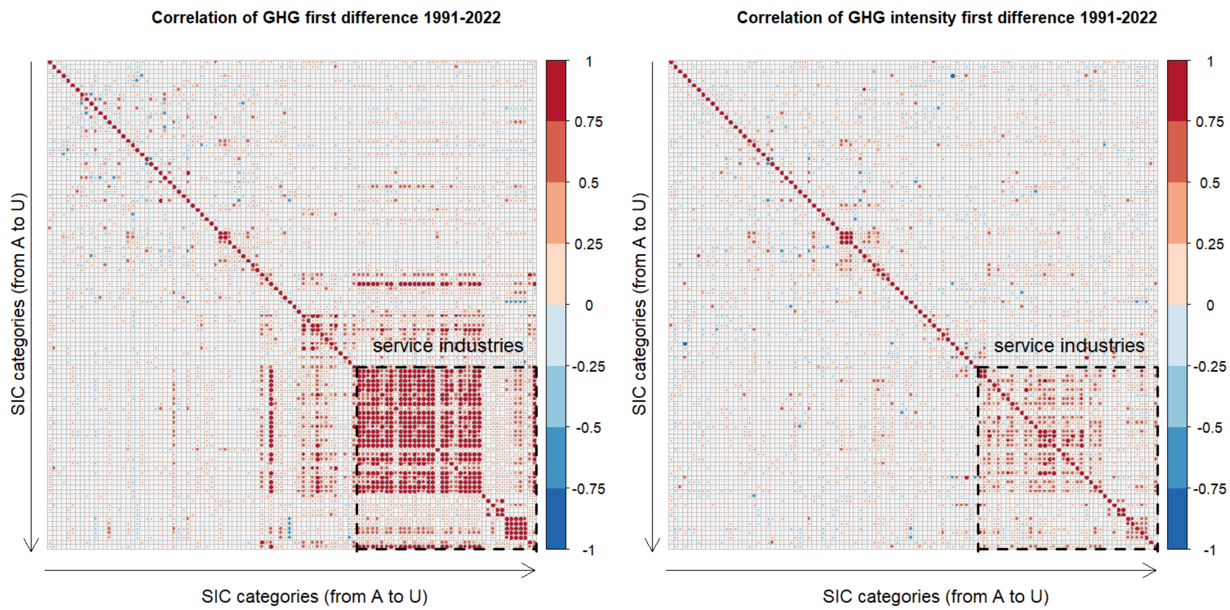

Figure 4: Correlation matrices for the first differences of emissions and emissions intensity.

energy-conversion efficiency gains and offshoring contribute significantly to observed trends.

Our objective, however, is not to develop a more granular decomposition. Instead, we extend this literature by examining the degree of co-movement in emissions intensity across industries. While our framework could, in principle, be combined with deeper decompositions – for example, to assess whether changes in conversion efficiency identified by Hardt et al. (2018) are driven by common or idiosyncratic factors – we focus here on quantifying the extent to which industry-level changes move together or diverge over time. Because the within effect accounts for most of the decline in aggregate emissions intensity, we concentrate on explaining its dynamics. Specifically, we decompose the within effect into components driven by industry-specific (idiosyncratic) shocks and those arising from common shocks. To implement this decomposition, it is essential to account for co-movement among groups of related industries, which we document in the next section.

2.4 *Fact 3*: There exists substantial co-movement at the block level

Figure 4 displays the correlation matrices of year-on-year changes in emissions (left panel) and emissions intensity (right panel) across industries from 1991 to 2022. Each cell shows the correlation between two SIC industries, with darker red indicating stronger positive correlation and darker blue indicating stronger negative correlation. Both matrices exhibit a clear block structure – clusters of industries with systematically higher correlations. This structure is particularly pronounced among service industries (outlined in the dashed box), while manufacturing and primary industries display weaker and more dispersed correlation patterns. In Supplementary Information E, we show the same results as in Fig. 4, but split into two periods, 1991–2006 and 2007–2022. The block structure

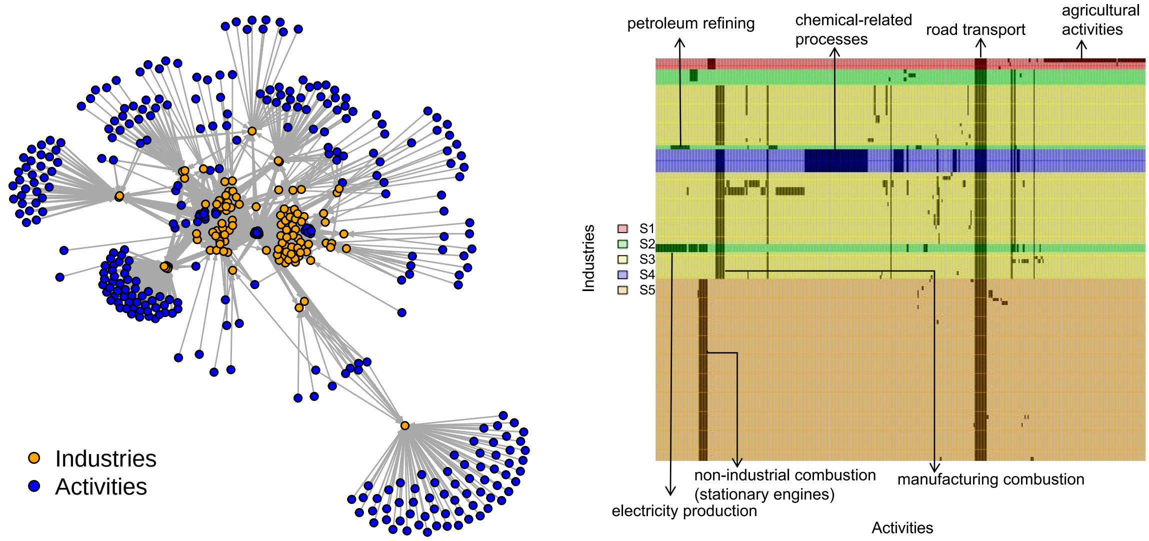

Figure 5: Visual representations of the correspondence table according to the Eurostat Manual for Air Emissions Accounts - 2012 edition (correspondence between SNAP97 - CRF/NFR and NACE Rev.2.). The left panel shows the bipartite network where an edge exists if an industry uses an activity. The right panel shows the incidence matrix, where a square is dark if the industry on y-axis uses the activity on the x-axis. The industries are ordered based on the initial ordering of SIC07. The activities are ordered based on SNAP code classification. The colors represent blocks chosen by the community detection algorithm (see Section 3.2.4), which we will use in our block-level factor model. All vertical lines spanning the entire matrix, showing activities that are performed by all industries, are related to road transportation.

becomes somewhat more pronounced in the latter period, suggesting an increasing synchronization of emissions intensity across subsets of industries.

The presence of such distinct clusters of co-movement suggests that emissions dynamics are not purely idiosyncratic or global, but rather mediated through shared exposure to emission-generating activities. This pattern is reminiscent of insights from the production network literature, which shows that the interdependence between industries can generate substantial co-movement between industry growth rates (Acemoglu et al., 2012), so much so that it becomes difficult to separate co-movement due to input-output interdependencies from co-movement due to common factors (Foerster et al., 2011). In our context, shared polluting activities, such as road transport, create co-movement between industries, but these may or may not be *global*. In the production network literature, Foerster et al. (2011) show that one can exploit data on the production network to filter out network-related correlations from raw growth rates, before applying a classic factor model. Here we do not have the detailed information needed to perform a similar exercise, but we will argue that grouping industries into blocks that share similar activities can achieve a similar purpose: by allowing for non-global co-movement, we can better identify purely idiosyncratic and purely global co-movement.

What explains this block structure? Emissions data are not measured directly at the industry level but, following UNFCCC conventions, are first compiled for key emitting activities and sources (e.g., urban driving by diesel cars) and then allocated to industries using emission factors (tons of CO_2 per unit of fuel). In the UK, Ricardo Energy and Environment (REE) maps these activity-level data from the National Atmospheric Emissions Inventory (NAEI) to Standard Industrial Classification (SIC) categories by multiplying the quantity of each activity by its corresponding emission factor (NAEI, 2022). This allocation is independent of time as it simply categorically defines the emission activities that are relevant to the business production of each industry.

Because many industries draw on common underlying activities – such as road transport or electricity generation – the resulting estimates naturally reflect clusters of correlated emissions dynamics. These correlations, while partly shaped by the allocation procedure, capture genuine shared exposure to polluting activities and motivate our decision to account explicitly for block-level co-movement in the subsequent analysis. Further details on the allocation process are provided in Supplementary Information A.1. Although the exact correspondence table used by REE is not publicly available, Fig. 5 illustrates a comparable allocation scheme based on the Eurostat Manual for Air Emissions Accounts (Eurostat, 2012)[4].

The left panel of Fig. 5 highlights three key features that are likely present in the UK data. First, many emission-generating activities span multiple industries. Second, some industries perform unique activities (e.g., NFR/CRF "Field burning of agricultural wastes" is linked only to NACE Rev2 "01 agriculture"). Third, some industries appear to rely on a common subset of activities. This means that industry-level emissions arise from both shared and industry-specific sources. The right panel of Fig. 5 shows the incidence matrix of the bipartite network between activities and industries. The x-axis lists distinct emission-generating activities, while the y-axis shows the industries involved. A black cell indicates that a given activity is used by a particular industry. Some activities, such as "diesel car driving", are widely shared across industries. Others are unique and connect only to specific sectors. Importantly, certain subsets of activities are linked to mutually exclusive subsets of industries, forming clear "blocks" of shared exposure.

The way emissions are measured and allocated – from activity-level sources to industries – naturally generates both shared and industry-specific components, helping to explain the global co-movement and block-level synchronization observed in Figure 4. Together with the other stylized facts, this motivates a formal accounting framework, which we introduce in the next section. We show that this framework, grounded in the way the data are constructed, provides a useful representation of the dynamics of aggregate and industry-level emissions intensity, and serves as the basis for our factor model decomposition.

[4] The manual, based on Eurostat's inventory-first approach (Eurostat, 2009), underpins the emissions accounting in our dataset (ONS, 2023), following UN System of Environmental Economic Accounts standards. See Supplementary Information A.1 for details.

3 Methods

The empirical patterns documented above, particularly the strong within-industry contribution to aggregate emissions reductions and the presence of distinct blocks of co-movement across industries, point to a deeper structure in the emissions data. These features suggest that emissions dynamics are shaped both by shared sources of emissions (such as common energy-related activities) and by industry-specific responses to regulation, technology, and changing production processes. To investigate these dynamics more formally, this section develops the conceptual and empirical foundation for the block-level factor model used in our analysis. Our objective here is to bridge the structure of the data, economic interpretation, and statistical modeling, providing a framework that is both theoretically grounded and empirically tractable.

We begin by outlining how emissions are constructed and attributed in the UK data: emissions are first calculated at the level of activities (e.g., combustion, transport, industrial processes) and then allocated to industries based on their involvement in these activities. This allocation process itself creates a network structure in which industries share exposure to certain emissions sources – leading to both global and block-level correlation in emissions intensity. We then introduce a simplified model of this allocation mechanism, expressing an industry's emissions intensity as the weighted sum of activity-level emissions intensities. This decomposition clarifies how co-movement can arise mechanically from the data structure, and provides the theoretical rationale for the block-level decomposition used in our empirical model. Finally, we show how this allocation-based logic maps naturally into a dynamic factor model, where common components reflect shared activity shocks and the idiosyncratic component captures industry-specific changes in emissions efficiency.

3.1 The allocation-based DGP

The precise details of the emissions allocation process are not publicly disclosed and likely involve considerable complexity. Nevertheless, the dataset must satisfy a straightforward accounting tautology: by construction, the industry-level emissions intensities reported in the data reflect the aggregation and allocation of emissions across underlying activities. We use this identity as the basis for a stylized data-generating process (DGP) that captures the essential structure of the dataset and serves as the foundation for our empirical framework.

3.1.1 From the emission allocation process to a statistical model

We base our approach on the definitions of "activities" and "emission factors" used by the NAEI and the ONS emission allocation process. In this context, activities are measured by the units of energy or fuel consumed in performing each activity (e.g., the volume of diesel used per mile traveled). Emission factors, in turn, represent the emissions generated per unit of energy or fuel consumed (e.g., the amount of CO_2 produced from the combustion of one unit of diesel). With this definition of activities and emission factors, we can write the total emissions of the industry i at time t as

$$E_{it} = \sum_j E_{ijt} = \sum_j \left[\frac{E_{ijt}}{A_{ijt}} \times \frac{A_{ijt}}{Y_{it}} \times Y_{it} \right], \tag{3}$$

where A_{ijt} is the amount of activity j carried out by industry i at time t, and Y_{it} is the output of industry i at time t. The first factor inside the sum in Eq. 3 represents the amount of emissions produced for each unit of activity j by industry i, and the second term represents the amount of activity j needed for producing each unit of GVA in industry i. Denoting these objects as θ_{ijt} and r_{ijt}, respectively, and noting that Y_{it} can be brought outside the sum because it is not related to activity j, we can express emissions intensity as

$$\frac{E_{it}}{Y_{it}} \equiv \mathcal{E}_{it} = \sum_j \theta_{ijt} r_{ijt}. \tag{4}$$

Eq. 4 demonstrates that industry-level emissions intensity can be decomposed into the sum of contributions from individual activities. The emissions generated by each unit of activity j are determined by the product of the emission factor θ_{ijt}, which quantifies the emissions produced per unit of activity (the activity emissions intensity, or "emission factor"), and the "activity requirement" r_{ijt}, which indicates how much of this activity is needed to produce one unit of GVA in industry i.

To go from mere accounting to a useful statistical framework, we introduce assumptions that we think are the most plausible.

1. The "emission factor" θ_{ijt} is the same in all industries, and remains constant over time,

$$\theta_{ijt} = \frac{E_{ijt}}{A_{ijt}} = \theta_j \quad \forall t, \forall i. \tag{5}$$

   This assumption is supported by our detailed analysis in Appendix A.2.3, which confirms that assuming constant emission factors is likely to be a good first-order approximation.

2. The "activity requirement" term, r_{ijt}, follows a random walk process[5],

$$r_{ij,t} = r_{ij,t-1} + \beta_{ij} u_{jt} + \epsilon_{ijt}, \tag{6}$$

   where, u_{jt} represents a shock to activities requirement which is common to all industries, β_{ij} represents the extent to which this common shock affects industry i, and ϵ_{ijt} represents a shock to activity requirement j that is specific to industry i. These shocks allow us to isolate industry-specific efforts to reduce the use of an activity, from reductions in the use of an activity that is common to all industries.

[5] We could have chosen a stationary process; in this case, this framework suggests estimating a dynamic factor model in levels, rather than in first differences. In view of Onatski and Wang (2021)'s results showing that estimating factor models in levels on non-stationary data leads to spurious results, we prefer to estimate our model in first differences. We could also have chosen a factor structure for $\Delta \log r_{ijt}$ instead of for Δr_{ijt}. While it would perhaps be more intuitive to think of common shocks as having a multiplicative impact on activity requirements, the additive model produces neater results, including for aggregation.

3. The two shocks are independent[6],

$$\text{Cov}(\epsilon_{ijt}, u_{jt}) = 0, \quad \forall i, \forall j, \tag{7}$$

and the idiosyncratic efforts to reduce activities are cross-sectionally independent,

$$\text{Cov}(\epsilon_{ijt}, \epsilon_{sjt}) = \text{Cov}(\epsilon_{ijt}, \epsilon_{smt}) = 0, \quad \forall i \neq s, \forall j \quad \text{and} \quad j \neq m. \tag{8}$$

Note that these assumptions do not exclude the possibility of cross-sectional dependence among activity requirements, a topic that will be explored further in the subsequent sections. Plugging Eq. 6 and 5 into Eq. 4 and taking first differences, we can express the change in emissions intensity in industry i at time t as

$$\Delta\mathcal{E}_{it} = \sum_j \theta_j \beta_{ij} u_{jt} + \sum_j \theta_j \epsilon_{ijt}. \tag{9}$$

The second sum in Eq. 9 represents the industry-specific change in emissions intensity. To ease notation, we introduce $\omega_{it} \equiv \sum_j \theta_j \epsilon_{ijt}$, so Eq. 9 becomes

$$\Delta\mathcal{E}_{it} = \sum_j \theta_j \beta_{ij} u_{jt} + \omega_{it}, \tag{10}$$

and where we note for future reference that our assumption on the cross-sectional independence of the shocks ϵ extends to the term ω,

$$\text{Cov}(\omega_{it}, \omega_{st}) = \text{Cov}\Big(\sum_j \theta_j \epsilon_{ijt}, \sum_j \theta_j \epsilon_{sjt}\Big) = \sum_m \sum_j \theta_m \theta_j \text{Cov}(\epsilon_{ijt}, \epsilon_{smt}) = 0, \quad \forall i \neq s. \tag{11}$$

Eq. 10 shows that the change in industrial emissions intensity can be decomposed into two distinct components: a component that arises due to common shocks to activity requirements (u_{jt}), and a component that arises from idiosyncratic changes in activities requirements (ϵ_{ijt}). Thus, thanks to the above assumptions, we can effectively link our framework to the factor model.

3.1.2 Link to a factor model

From Eq. 10, and using the result (11), we can calculate the covariance between emission reductions in two industries,

$$\text{Cov}(\Delta\mathcal{E}_{it}, \Delta\mathcal{E}_{st}) = \underbrace{\sum_{j \neq m} \theta_j \theta_m \beta_{ij} \beta_{sm} \text{Cov}(u_{jt}, u_{mt})}_{\text{contrib. of cov. between activity requirements}} + \underbrace{\sum_j \theta_j^2 \beta_{ij} \beta_{sj} \text{Var}(u_{jt})}_{\text{contrib. of common activities}}, \ \forall i \neq s. \tag{12}$$

[6] More precisely, to be complete, we assume the activity j innovation ϵ_{ijt} is orthogonal not only to the common shocks for activity j, but also orthogonal to common shocks to all other activities j'. Formally, $\text{Cov}(\epsilon_{ijt}, u_{j't}) = 0$ for all j'.

The overall covariance can be expressed as the sum of two different components. The first term corresponds to a co-movement due to the fact that activity requirements themselves are co-moving. It shows that two industries can exhibit co-movement not only through shared activities but also because the requirements of their respective activities tend to move together. For example, non-metallic minerals industries and agriculture industries rarely share common activities, yet they can co-move due to a national policy related to, say, the use of coke. This can occur because non-metallic minerals industries have an emission activity involving the use of coke for lime production, while agriculture has an emission activity related to stationary combustion using coke. The second term, instead, represents a co-movement due to the fact that two industries share a specific activity.

In a standard exact factor model, the covariance between two industries is entirely attributed to the common factors, with no correlation allowed between their idiosyncratic components. By contrast, an approximate factor model allows for some residual cross-sectional correlation in the idiosyncratic shocks. Applying this to the DGP above, we can infer that the common factors estimated from a factor model would reflect the two sources of co-movement identified in Eq. 12: (i) co-movement arising because industries share common activities, and (ii) co-movement arising because the activities themselves are correlated, even when industries do not share them directly. This interpretation parallels the insights of Foerster et al. (2011), who show that in production network settings, a naively estimated factor model on industry growth rates often captures co-movement generated by input-output linkages, rather than a true common shock affecting all industries.

To establish more formally a connection between the DGP that we derived from the allocation process, Eq. 10, and the DGP from a factor model, consider a dynamic factor model with K common factors [7]

$$\Delta \mathcal{E}_{it} = \sum_{k=1}^{K} B_{ki} F_{kt} + \nu_{it} \tag{13}$$

where ν_{it} are idiosyncratic components, F_{kt} are the factors and B_{ki} are the loadings. By comparing to Eq. 10, we could assume that the idiosyncratic components of both equations are the same, $\nu_{it} = \omega_{it}$, so the part corresponding to common shocks to activities requirements would be identified by the factors in the factor model,

$$\sum_{k=1}^{K} B_{ki} F_{kt} = \sum_{j} \theta_j \beta_{ij} u_{jt}. \tag{14}$$

This mapping between the allocation-based DGP and the factor model (Eq. 14) does not hold in full generality. However, as we discuss in Supplementary Information D, it is exact in two common cases: (i) when industry emissions are driven by a single activity, and (ii) when the impact of common shocks on activity requirements is homogeneous across activities within each industry ($\beta_{ij} = \beta_i$ for all j). These cases illustrate the conditions under which the factor model can be viewed as an exact representation of the allocation-based DGP. In more general settings, the factor model remains a useful approximation, capturing the dominant sources of co-movement implied by Eq. 10.

[7] To avoid notational clutter, here we omit the complication due to the fact that in practice the data is standardized. See Supplementary Information.

To sum up, industry-level emissions are constructed from the emissions of the activities in which industries are engaged. Because industries often share activities, their emissions naturally co-move. Additional co-movement can arise from correlations across activity requirements and, potentially, from residual dependencies among idiosyncratic shocks (which we have not modeled explicitly). While there is no perfect one-to-one mapping from this allocation-based data-generating process to a factor model[8], the framework we propose can accommodate the main sources of co-movement implied by the accounting tautology. In doing so, it provides a coherent interpretation of the common components estimated by our factor model and clarifies what is captured as global, block-level, and idiosyncratic variation in industry-level emissions intensities.

3.2 Block-level factor model

3.2.1 Model

We use an approximate dynamic factor model following the approach in Chatzivgeri et al. (2019), that allows for mild cross-sectional dependence (see also, Mumtaz and Theodoridis (2017), Mumtaz and Musso (2021)), global common factors, and block-level common factors[9]. We collect the series of our panel into a $N \times T$ matrix X with elements X_{it} denoting the (normalized) change of emissions intensity of industry i at time t, defined as

$$X_{it} \equiv \frac{\Delta\mathcal{E}_{it} - \mu_i}{\sigma_i}, \tag{15}$$

where μ_i and σ_i denote the population mean and standard deviation of $\Delta\mathcal{E}_{it}$ for industry i. In the empirical implementation, we estimate these parameters using the sample mean and standard deviation of each industry's time series.[10]

We assume that each X_{it} can be decomposed into three components, namely a global factor, block-specific factors, and a purely idiosyncratic component, that is

$$X_{it} = B_i^C F_t^C + B_i^S F_t^S + \nu_{it}, \tag{16}$$

where F_t^C denotes a set of K^C factors, with associated factor loadings B_i^C, that are common across all industries, F_t^S denotes a set of K^S block-level factors, with associated factor loadings B_i^S, and ν_{it}

[8] If detailed data on the mapping between industries and activities (the full weighted bipartite graph) were available, one could design an identification strategy for the model. Unfortunately, this information is not disclosed for the UK, and while we have Eurostat-based mappings, these are unweighted.

[9] An alternative approach would have been to use sparse PCA (Despois and Doz, 2023, see e.g.), which adds to the standard PCA-estimated factor model a regularisation term for the loadings. This generates loadings that are mostly zero, so that each factor is effectively associated with an endogenously determined set of units. Here, we prefer to define the blocks ourselves, and use our Bayesian model that also allows for autocorrelation and mild cross-sectional variation.

[10] While removing the mean is absolutely necessary to estimate a factor model, normalising by the standard deviation is optional here, as all our time series are in the same units. In practice, we have found that estimating a factor model on the non-normalised time series gives poor results, with a highly volatile common factor associated with substantial loadings from only a handful of industries. This is not surprising given the heterogeneity of volatilities seen in Fig. 2.

denotes the idiosyncratic components. While F_t^C denotes the "full economy" factors for emissions intensity, F_t^S denotes factor with loadings that are set equal to zero whenever industry i does not belong to the block s, where $s = 1, \ldots, S$.

Each of these unobserved factors and the idiosyncratic components are assumed to follow independent $AR(p)$ processes with normal errors. For ease of notation we collect these factors in the $T \times (K^C + N \times K^S)$ matrix F, where N is the number of industries in the model. The k^{th} column of this matrix is described by the process

$$F_{kt} = c_k + \sum_{l=1}^{p} b_{k,l} F_{k,t-l} + \sigma_k^{1/2} e_{kt}, \quad e_{kt} \sim N(0,1), \tag{17}$$

and similarly, the N idiosyncratic components are described by the $AR(q)$ process

$$\nu_{it} = \sum_{l=1}^{q} d_{i,l} \nu_{i,t-l} + h_i^{1/2} e_{it}, \quad e_{it} \sim N(0,1). \tag{18}$$

Notice that in order for the model to remain parsimonious and ensure identification, we have assumed that the block factors are uncorrelated with each other. This assumption is common in the literature, but it could be relaxed by imposing alternative identifying restrictions along the lines of Bai and Wang (2015). We do not pursue this route in this paper because our dataset is limited.

3.2.2 Estimation

Before estimating the model in first differences, we examine the time-series properties of emissions intensity. Augmented Dickey–Fuller (ADF) tests, which take the presence of a unit root as the null hypothesis, reject the null for only 22 out of 106 industries when applied to the series in levels. In contrast, when applied to first differences, the ADF tests reject the unit-root null for all 106 industries. As a complementary exercise, we apply KPSS tests, which take stationarity as the null hypothesis. The KPSS tests reject stationarity in levels for 94 out of 106 industries, while in first differences they reject stationarity for only 12 industries. Taken together, the evidence from both tests indicates that emissions intensity is well approximated as an integrated process of order one, I(1). This finding supports our decision to estimate the factor model in first differences, thereby focusing on changes in emissions intensity and reducing concerns about spurious common factors arising from non-stationary data.

We estimate this dynamic factor model by Bayesian methods, using a Gibbs sampling algorithm as in McCulloch and Tsay (1994), which is fairly standard and described in detail in the appendix of Chatzivgeri et al. (2019). In short, the algorithm is based on successively sampling from the conditional posterior distribution of F_t and the remaining parameters. It exploits the fact that given a draw for F_t, the model collapses to a series of linear regressions with standard posterior distributions. As our data is sampled at the annual frequency we fix the lag lengths at $p = q$=1. The choice of the number of common and country specific factors is a key specification issue. Given our small sample size, we select one factor at both global and block level (i.e., $K^C = K^S = 1$).

3.2.3 Variance decomposition

Given Eq. 16 we can decompose the total variance of emissions intensity as

$$\mathrm{Var}(X_{it}) = (B_i^C)^2 \mathrm{Var}(F_t^C) + (B_i^S)^2 \mathrm{Var}(F_t^S) + \mathrm{Var}(\nu_{it}). \tag{19}$$

This equation can be used to estimate the contribution of the global- and block-level common factors to the variance of each series. The model is subject to the usual scale and identification issues affecting factor models. First, the sign of the factors and the factor loadings is not identified separately. Secondly, since the scale of the factors is unidentified we impose the variance of the factors to be equal to one, using

$$\tilde{F}_t^C = \frac{F_t^C}{\sqrt{\mathrm{Var}(F_t^C)}}, \quad \tilde{F}_t^S = \frac{F_t^S}{\sqrt{\mathrm{Var}(F_t^S)}},$$

and rescale the loadings accordingly,

$$\tilde{B}^C = B^C \times \sqrt{\mathrm{Var}(F_t^C)}, \quad \tilde{B}^S = B^S \times \sqrt{\mathrm{Var}(F_t^S)}.$$

We can then rewrite Eq. 19 as

$$\mathrm{Var}(X_{it}) = (\tilde{B}_i^C)^2 + (\tilde{B}_i^S)^2 + \mathrm{Var}(\nu_{it}). \tag{20}$$

Then, the contributions of the global- and block-level common factors are:

$$\begin{aligned} V^C &= \frac{(\tilde{B}_i^C)^2}{(\tilde{B}_i^C)^2 + (\tilde{B}_i^S)^2 + \mathrm{Var}(\nu_{it})}, \\ V^S &= \frac{(\tilde{B}_i^S)^2}{(\tilde{B}_i^C)^2 + (\tilde{B}_i^S)^2 + \mathrm{Var}(\nu_{it})}. \end{aligned} \tag{21}$$

Equation 21 shows how the contribution of the common factor to the variance of an industry (and thus to aggregate volatility of emissions intensity) is determined by the loadings. Estimates of V^C and V^S can be used to infer if the variance of emissions has been driven by block-level factors or events that are common across the economy. Similarly a comparison of the *ith* data series X_{it} with the "fitted values" $\hat{X}_{it}^C = B_i^C F_t^C$ and $\hat{X}_{it}^{C+S} = B_i^C F_t^C + B_i^S F_t^S$ can be used to assess how the factors contribute to the changes in emissions over time.

3.2.4 Classifying industries into blocks

To implement our factor model in a way that respects the accounting structure of the data, we first classify industries into blocks based on their shared exposure to emission-generating activities. Using the Eurostat-based mapping introduced in Figure 5, we apply community detection methods to identify groups of industries that rely on many of the same activities. This bookkeeping step is crucial: by capturing clusters of industries with shared sources of emissions, it allows the factor

Block	General Description of Industries
S1	Agriculture and related industries
S2	Oil, gas, electricity, and petroleum production
S3	General manufacturing, waste management, and construction
S4	Chemical-related manufacturing
S5	Transportation and service industries

Table 2: Classification of industries into blocks. This classification is based on a community detection algorithm on the network where nodes are industries and links represent how many activities two industries have in common. It is different from broader industry groupings. We have labeled each block manually. For a comprehensive list of industries included in each block, see Table 7 in SI A.

model to distinguish between truly global shocks, block-level shocks, and industry-specific dynamics. We use the Louvain algorithm, which includes a resolution parameter enabling us to obtain a relatively small number of blocks while letting the data determine their exact composition. This procedure yields an intuitive partition of industries into five blocks. Detailed descriptions of these blocks and the classification process are provided in Figure B.1 in Supplementary Information.

Fig. 5, right, shows that the identified communities often correspond clearly to shared activities. The Figure also shows that the blocks do not perfectly correspond to a division by upper layers of the SIC hierarchy, showing that this classification scheme is non-trivial. Table 7 in Supplementary Information provides the full details of the allocation of industries to blocks. In Table 2 below, we provide intuitive names for the blocks, based on manual inspection.

4 Results

4.1 Contributions to the industry-level variance in emissions intensity

We first calculate the share of the total variance in changes in emissions intensity explained by the global and block-level factors for each industry (Figure 6). The figure displays, for each industry (ordered by the five blocks introduced in Section 3), the proportion of variance explained by the global factor (blue), block-level factors (red), and the idiosyncratic component (gray). Industries are grouped by block, with dashed lines separating the five clusters. For readability, industry labels along the horizontal axis are abbreviated; the full correspondence to SIC categories is provided in Table 7 in the Appendix.[11]

Our variance decomposition reveals that the global common factor makes limited contributions to the variance of the changes in emissions intensities of the industries, except for transportation and service industries (S5). This suggests that the influence of truly economy-wide forces on emissions

[11] This table lists all industry labels used along the horizontal axis, as well as the full SIC sector names.

intensity is limited in most sectors, but present in services and transport, likely reflecting shared exposure to upstream decarbonization of inputs like road transport. In other words, an explanation for the pattern in Figure 6 is that for service industries, most of their emissions are related to universal inputs, so a high share of their volatility is explained by a common factor. In contrast, the

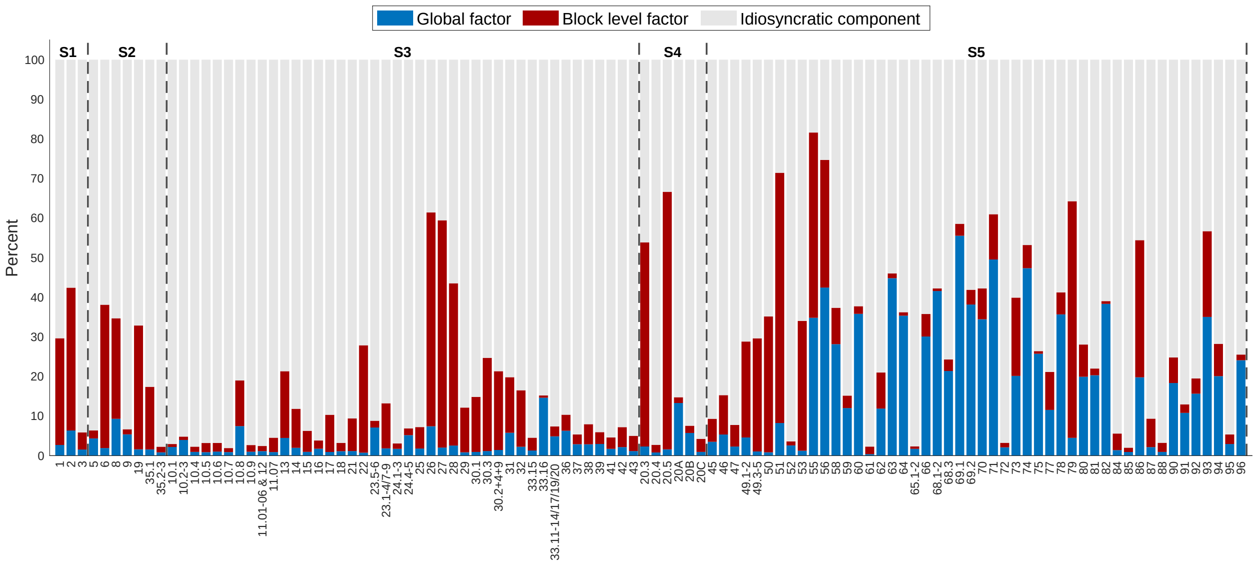

Figure 6: Variance explained by the block-level factor model. The figure highlights the explanatory power of the block-level model in capturing the underlying industrial structure.

block-level factors are highly influential in a number of sectors, especially in manufacturing (S3) and chemical industries (S4). For example, sectors 26–29 (electrical equipment, machinery, vehicles) and chemical subsectors (C20.3–C20.5) show variance shares ranging from 20% to over 60% explained by their respective block-level components. Block S2, which includes oil, gas, electricity, and refined petroleum, also shows very strong block-level explanatory power. The presence of strong block-level co-movement in these clusters indicates structural similarities in production and emissions, suggesting that certain industries are jointly exposed to shocks affecting shared emission-generating activities, such as for instance chemical inputs, industrial heat, or logistics. The idiosyncratic component still plays a major role in sectors like apparel, construction, manufacturing and some services, where emissions dynamics likely depend on firm-level characteristics. The dominance of the idiosyncratic component in these sectors suggests that emissions reductions are largely the result of sector-specific strategies or technologies, rather than broad-based macroeconomic trends.

In Supplementary Information C, we present the estimated global and block-specific factors, along with their posterior distributions. The cumulative global factor estimate exhibits a consistent downward trend, with a particularly steep decline observed around 2005. The block-level factors for S3 (Manufacturing) and S5 (Services) also show negative values, suggesting that these block-level factors contribute to a reduction in emissions intensity over time. However, it is important to note that the factors for S1 (Agriculture), S2 (Fossil) and S4 (Chemicals) are not statistically significant, indicating a limited independent impact on the overall variation in emission intensities within these blocks.

In Supplementary Information C, we also provide a comparison between the results of the block-level model to those from a standard factor model. Overall we find that the block-level model, by providing additional structure, reduces the portion of the variation attributed to the idiosyncratic component, offering a better overall fit. The improved fit of the block-level model further underscores that emissions intensity dynamics are not driven solely by aggregate or purely idiosyncratic shocks, but also by structured co-movement within specific parts of the production network. Accounting for these block-level interdependencies is therefore essential for understanding how emissions reductions propagate across the economy.

A common concern with factor models is that they do not allow for time-varying loadings. One could think that, with climate change policies becoming more prominent, perhaps at different rates across industries, the variance explained by the common factor may have increased over time. Our data is too short to check for this rigorously, but in Supplementary Information E we divide the full sample into two periods and re-estimate the block-level factor model for each period to examine how the influence of common factors has evolved over time.

Figure E.2 reports the results for the full sample, while Figure E.3 repeats the exercise excluding the COVID-19 period. Overall, we do not find strong evidence that the variance explained by the global or block-level factors has changed over time. The results are broadly similar across the two specifications, with only a modest increase in the importance of the block-level factor for the services sector. This suggests that our findings are not driven by pandemic-specific dynamics.

4.2 Contribution to the variance of the aggregate emissions intensity

While the preceding analysis decomposes the sources of variation in emissions intensity at the industry level, what ultimately matters for assessing progress toward climate goals is the behavior of *aggregate* emissions intensity. As shown in Section 2.3, most of the aggregate change is explained by the within-industry component, rather than shifts in industry shares. We therefore focus here on decomposing this within component into the contributions of global, block-level, and idiosyncratic factors. This allows us to quantify the extent to which aggregate emissions intensity reductions are driven by economy-wide forces, coordinated shifts within groups of industries, or purely industry-specific dynamics.

Instead of decomposing the variance of aggregate emissions intensity directly, we focus on the variance of the within effect. This focus is justified because, as shown in Eq. 1, changes in aggregate emissions intensity can be expressed as

$$\Delta\mathcal{E}_t = \underbrace{\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it}}_{\text{“within effect”}} + \underbrace{\sum_i \bar{\mathcal{E}}_{it}\Delta w_{it}}_{\text{“between effect”}} \; , \tag{22}$$

where the first term captures changes in industry-level emissions intensities holding value-added shares constant, and the second term captures changes due to shifts in industry shares. As documented in Section 2.3, the within effect accounts for the vast majority of the aggregate change, making it the primary driver of aggregate decarbonization. Table 3 reports this decomposition for the long difference

$$\Delta\mathcal{E} = \mathcal{E}_{2022} - \mathcal{E}_{1990},$$

highlighted in the row "overall." To extend this analysis to volatility, we decompose the variance of the first difference of aggregate emissions intensity as

$$\begin{aligned}\mathrm{Var}(\Delta\mathcal{E}_t) = &\mathrm{Var}\Big(\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it}\Big) + \mathrm{Var}\Big(\sum_i \bar{\mathcal{E}}_{it}\Delta w_{it}\Big) \\ &+ 2\ \mathrm{Cov}\Big(\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it},\ \sum_i \bar{\mathcal{E}}_{it}\Delta w_{it}\Big).\end{aligned} \tag{23}$$

	Within	Between	Covariance
Overall	0.73	0.27	
Mean	0.65	0.35	
Variance	4.41	2.93	-6.33

Table 3: Relative contributions of the within and between effects to the mean and variance of the first difference in aggregate emissions intensity, and to the "overall" decline, that is between 1990 and 2022.

Table 3 reports the relative contributions of the within and between components to aggregate emissions intensity, both for the long difference between 1990 and 2022 (row "Overall") and for the mean and variance of its first difference. The within component explains roughly three-quarters of the overall decline and two-thirds of the mean change in emissions intensity, while its contribution to the variance is substantially larger than that of the between component. Given its dominant role in both the level and volatility of aggregate emissions intensity, the analysis that follows focuses on decomposing the within component into global, block, and idiosyncratic factors. To reveal how different factors contribute to the variance of the within effect (defined in Eq. 2) we can write

$$\begin{aligned}\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it} &= \sum_i \bar{w}_{it}(\sigma_i X_{it} + \mu_i) \\ &= \sum_i \bar{w}_{it}(\sigma_i(\tilde{B}_i^C\tilde{F}_t^C + \tilde{B}_i^S\tilde{F}_t^S + \nu_{it}) + \mu_i) \\ &= \sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^C\tilde{F}_t^C + \sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^S\tilde{F}_t^S + \sum_i \bar{w}_{it}\sigma_i\nu_{it} + \sum_i \bar{w}_{it}\mu_i,\end{aligned} \tag{24}$$

where the intermediate steps use Eqs. 15 and 20.

Figure 7 (left) decomposes changes in aggregate emissions intensity (through the within effect) into four components: global factors, block-level factors, idiosyncratic shocks, and a drift term. The drift term represents the long-run trend decline in emissions intensity, which averages roughly –0.006 per year. This figure is consistent with the aggregate trend shown in Figure 1: emissions intensity falls from 0.46 to 0.18 kton per £m of GVA between 1990 and 2022, a total reduction of 0.27. As shown in Table 3, 73% of this change is attributable to the within component, corresponding to

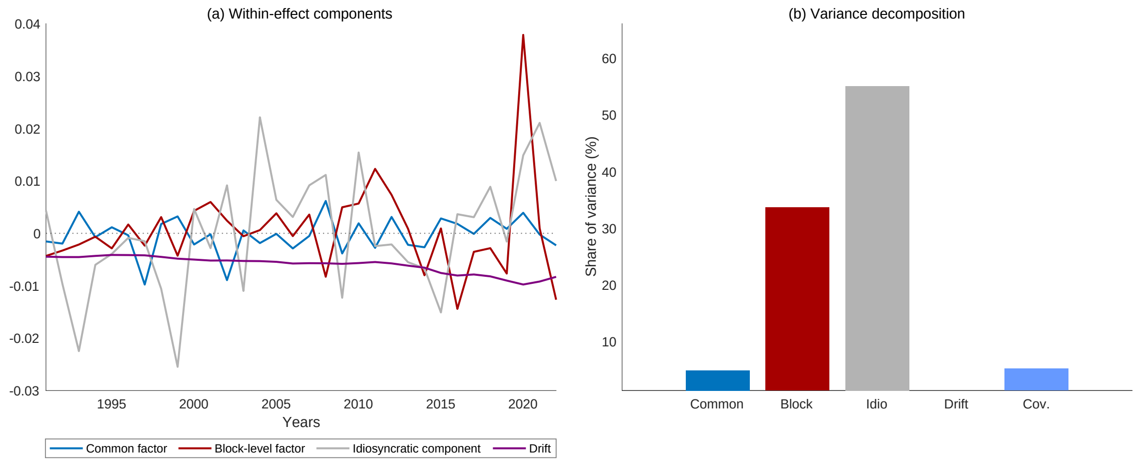

Figure 7: Analysis of contributions to the within effect. The left panel illustrates the time series of contributions from the common factor, idiosyncratic component, and drift to the within effect of the emissions data. These contributions are estimated with the block-level factor model. The right panel presents a stacked bar chart of the variance contributions.

a decline of about 0.19 over 33 years, or –0.006 per year. The figure also highlights the sources of short-run variation in aggregate emissions intensity. The idiosyncratic component (gray dashed line) is the most volatile over time, indicating that year-to-year fluctuations are primarily driven by industry-specific shocks rather than systematic factors. The global and block-level components (blue and red lines, respectively) exhibit less variability and remain near zero on average, while the drift term accounts for the steady long-term decline. Because the idiosyncratic component (the gray dashed line) is much more volatile than the contribution from the common factor, it will play a substantial role in explaining variations in emissions intensity. To test this, from Eq. 24, we compute the variance of the within effect as

$$\begin{aligned}\mathrm{Var}\Big(\sum_i \bar{w}_{it}\Delta\mathcal{E}_{it}\Big) &= \mathrm{Var}\Big(\sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^C\tilde{F}_t^C + \sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^S\tilde{F}_t^S + \sum_i \bar{w}_{it}\sigma_i\nu_{it} + \sum_i \bar{w}_{it}\mu_i\Big) \\ &= \mathrm{Var}\Big(\sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^C\tilde{F}_t^C\Big) + \mathrm{Var}\Big(\sum_i \bar{w}_{it}\sigma_i\tilde{B}_i^S\tilde{F}_t^S\Big) \\ &\quad + \mathrm{Var}\Big(\sum_i \bar{w}_{it}\sigma_i\nu_{it}\Big) + \mathrm{Var}\Big(\sum_i \bar{w}_{it}\mu_i\Big) + \mathrm{Cov}().\end{aligned} \tag{25}$$

The right panel of Figure 7 shows the variance decomposition of the within effect, as defined in Eq. 25. Idiosyncratic shocks account for the largest share of the variance, followed by a moderate contribution from block-level factors. The global factor plays only a minor role, while the drift and covariance correction terms are negligible. This pattern contrasts with Figure 6, where block-level factors explained a sizable portion of variance at the industry level. The apparent tension between these results highlights a key insight: although block-level co-movement is important for understanding micro-level variation across industries, its contribution largely washes out in the aggregate. This attenuation reflects differences in industry sizes and exposures, as well as the potential for offsetting shocks across sectors when aggregated.

Taken together, the two panels of Figure 7 underscore that aggregate emissions intensity dynamics are driven primarily by granular, industry-specific fluctuations. Macro-level decarbonization trends thus emerge not from pervasive common factors, but from the aggregation of diverse industry-level paths. Capturing these dynamics requires careful modeling of heterogeneity and aggregation, rather than relying exclusively on global or group-wide shocks.

5 Discussion

Panel data on industry-level emissions is a relatively recent development, and this study is among the first to systematically analyze these data. By documenting key empirical patterns – including the dominance of within-industry changes, the presence of distinct blocks of co-movement, and the predominance of granular shocks at the aggregate level – we have shown how these stylized facts motivate, and ultimately justify, the structure of our methodological approach. Our analysis highlights how the accounting foundations of the data and the observed patterns of co-movement can be reconciled within a block-level factor model, providing new insights into the drivers of aggregate decarbonization dynamics.

On the one hand, aggregate emissions reductions are driven primarily by within-industry improvements, underscoring the need to model heterogeneous, industry-specific dynamics. On the other hand, emissions intensities display substantial co-movement across industries, driven partly by economy-wide forces and partly by clusters of industries linked through shared, emissions-generating activities. These patterns reflect a layered structure: industry-specific behaviors, technologies, and policies overlaid on a complex, partly unobserved network of shared inputs and processes.

Capturing this structure requires balancing realism with tractability. A single global factor model would overlook important group-level dynamics and exaggerate the common component. Fixed-effects approaches, while straightforward, cannot capture latent structure or heterogeneous shock exposures. Fully structural, network-based models could, in principle, trace emissions through the complete bipartite mapping of industries to emitting activities, but these require detailed, country-specific weights over time – data not publicly available for the UK. Instead, we leverage Eurostat's emissions allocation structure as a stylized approximation and build a block-level dynamic factor model. This approach respects the layered nature of co-movement, accommodates flexible exposures, and produces a variance decomposition aligned with economically interpretable categories, offering a tractable yet insightful framework.

These findings are based on a 33-year period, during which the structure of emissions dynamics may itself have evolved. With the available data, models with time-varying loadings or evolving block structures are too difficult to estimate robustly. When we split the sample into two subperiods (1990–2006 and 2006–2022), we do not find strong, obvious evidence that the global or block components become more prominent in the latter period, but this remains an interesting hypothesis for future work.

We use a factor model in first-differences, which loses long-run information. This is a conservative choice, as a naive model in levels would risk capturing a spurious trend (Onatski and Wang, 2021). It would be interesting to revisit the results using models for the levels that can handle non-stationary and/or drifting common factors, such as Peña and Poncela (2006), Barigozzi et al. (2021) and Kaufmann and Strachan (2024). Intuitively, when units in levels follow a common factor integrated of order 1, they are themselves integrated of order 1, but because they follow the common factor, they are *co*-integrated. In this perspective, it is possible, or perhaps likely, that the share of variance (of the *levels*) explained by common factors will be substantially higher than what we find here, where we limit ourselves to "business-cycle-like" common factors, which affect all units' contemporaneous *changes* of emissions intensity, and where we measure the share of explained *volatility*. Another interesting possibility would be to use the "low-frequency econometrics" approach of Müller and Watson (2018), which explicitly separates low and high frequency movements to better estimate how long-term trends co-move.

Although our analysis focuses on the UK, the framework we develop is designed to be applicable across contexts. We use UK data because it is uniquely detailed: to our knowledge, it is the only long-run panel that links industry-level GHG emissions to economic output across a broad set of industries. While such data are not yet widely available for other countries, many economies are pursuing sector-specific decarbonization pathways within broader climate strategies. Our decomposition approach offers a tractable method for disentangling the structure of emissions dynamics elsewhere and could enable systematic cross-country comparisons as similar data become available.

Our results are also limited to the specific variables that we chose. It is possible that using gross output instead of gross value added as a denominator would give different results. Similarly, we consider only scope 1 emissions, while it is possible that some intensity reduction comes from offshoring, an issue we cannot speak to here.

Our results raise important questions for future work. We do not fully identify the source of the common factors driving emissions; future research could explore these structural origins – whether macroeconomic, regulatory, or technological – at both the national and industry levels. A particularly promising avenue would be to estimate similar models at the firm level, where variation in innovation, energy sourcing, and responsiveness to policy is likely to be even more pronounced.

6 Conclusion

This paper applies a block-level dynamic factor model to long-run UK industry-level emissions data to assess the relative contributions of industry-specific and common shocks to changes in aggregate emissions intensity. To our knowledge, this is the first application of dynamic factor analysis in this context, enabling a systematic decomposition of co-movement in emissions across industries. We use a newly assembled panel from the UK Office for National Statistics (1990–2022) that links greenhouse gas emissions to gross value added across 106 industries. While prior work has explored sectoral heterogeneity in emissions, the question of how industry-level dynamics shape aggregate decarbonization remains largely unresolved, particularly in terms of identifying where reductions originate and how they aggregate.

Our first result confirms that declines in aggregate emissions intensity are primarily driven by improvements within industries, rather than shifts in economic activity across industries. Emissions reductions have largely been achieved through industry-specific improvements, such as technological upgrading, process innovation, or the use of less polluting inputs, rather than through structural reallocation away from more polluting sectors.

However, while most of the aggregate change is due to industry-level change, this does not mean that most of the aggregate change is due to industry-level efforts. It is possible that industry-level changes are synchronous, which would suggest that industry-level change is actually the result of global forces. To investigate this, we used a factor model.

It reveals that idiosyncratic, industry-specific shocks do account for the majority of variation in emissions intensity, contributing much more than the global and block-level factors to the within-industry component of aggregate fluctuations. This indicates that emissions intensities are indeed primarily shaped by industry-specific factors, such as industry-specific technological progress and regulatory shifts. At the same time, while it does not matter much in the aggregate, we identify a meaningful economy-wide component that captures co-movement across all industries, and is particularly salient in service industries. This can be the result of aggregate policy or macroeconomic forces, or, as we discovered by investigating how the data is constructed, it can reflect a lower use of polluting activities typically carried on by all industries, such as road transport.

Together, these findings suggest that past decarbonization has been mostly the result of industry-specific improvements, although heterogeneous responses to common drivers of change have also played a noticeable role. To the extent that aggregate decarbonization has been driven by policy, this suggests that it is industry-specific policies that have been more effective, so far. Our framework offers a generalizable tool for disentangling these forces and provides a foundation for future work that bridges data, economic structure, and statistical modeling in the analysis of emissions dynamics.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.

Agarwala, M. and Martin, J. (2022). Environmentally-adjusted productivity measures for the UK. Technical report, The Productivity Institute Working Paper.

Amadei, C., Dosi, C., and Pintus, F. J. (2024). Energy intensity and structural changes: Does offshoring matter? *Fondazione Eni Enrico Mattei Working paper*, 2024-26.

Ang, B. W. and Zhang, F. Q. (2000). A survey of index decomposition analysis in energy and environmental studies. *Energy*, 25(12):1149–1176.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bai, J. and Wang, P. (2015). Identification and bayesian estimation of dynamic factor models. *Journal of Business & Economic Statistics*, 33(2):221–240.

Barigozzi, M., Cuzzola, A., Grazzi, M., and Moschella, D. (2024). Factoring in the micro: A transaction-level dynamic factor approach to the decomposition of export volatility. *Oxford Bulletin of Economics and Statistics*.

Barigozzi, M., Lippi, M., and Luciani, M. (2021). Large-dimensional dynamic factor models: Estimation of impulse–response functions with i (1) cointegrated factors. *Journal of Econometrics*, 221(2):455–482.

Bańbura, M., Giannone, D., and Reichlin, L. (2011). Nowcasting. In Clements, M. P. and Hendry, D. F., editors, *The Oxford Handbook of Economic Forecasting*. Oxford University Press, Oxford.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

Chatzivgeri, E., Mumtaz, H., Tavasci, D., and Ventimiglia, L. (2019). Common and country specific factors in the distribution of real wages. *Economics Letters*, 184:108582.

Colmer, J., Martin, R., Muûls, M., and Wagner, U. J. (2024). Does pricing carbon mitigate climate change? firm-level evidence from the European Union emissions trading system. *Review of Economic Studies*. Forthcoming.

Delle Chiaie, S., Ferrara, L., and Giannone, D. (2022). Common factors of commodity prices. *Journal of Applied Econometrics*, 37(3):461–476.

Despois, T. and Doz, C. (2023). Identifying and interpreting the factors in factor models via sparsity: Different approaches. *Journal of Applied Econometrics*, 38(4):533–555.

Doda, B. (2014). Evidence on business cycles and CO2 emissions. *Journal of Macroeconomics*, 40:214–227.

Eurostat (2009). Manual for air emissions accounts. Available online at `https://ec.europa.eu/eurostat/documents/3859598/5910405/KS-RA-09-004-EN.PDF.pdf/55f22c82-7935-43b0-8c95-30b352c559b3?t=1414781559000`.

Eurostat (2012). Annex 1 to the manual for air emissions accounts: Correspondence between SNAP97 (CRF/NFR) and NACE Rev.2. Available online at `https://ec.europa.eu/eurostat/documents/1798247/6191529/Annex1-AEA-Manual-(SNAP-CRFNFR-NACE2)-v20.xls/91bfd3ee-58e7-4265-a7a7-f1a3ea73b73e`.

Fankhauser, S., Smith, S. M., Allen, M., Axelsson, K., Hale, T., Hepburn, C., Kendall, J. M., Khosla, R., Lezaun, J., Mitchell-Larson, E., et al. (2022). The meaning of net zero and how to get it right. *Nature Climate Change*, 12(1):15–21.

Foerster, A. T., Sarte, P.-D. G., and Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy*, 119(1):1–38.

Forni, M. and Reichlin, L. (1998). Let's get real: a factor analytical approach to disaggregated business cycle dynamics. *The Review of Economic Studies*, 65(3):453–473.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

GFANZ (2022). Financial institution net-zero transition plans. Available online at `https://assets.bbhub.io/company/sites/63/2022/10/Financial-Institutions-Net-zero-Transition-Plan-Executive-Summary.pdf`.

Griliches, Z. and Regev, H. (1992). Productivity and form turnover in Israeli industry: 1979-1988.

Hardt, L., Owen, A., Brockway, P., Heun, M. K., Barrett, J., Taylor, P. G., and Foxon, T. J. (2018). Untangling the drivers of energy reduction in the uk productive sectors: Efficiency or offshoring? *Applied Energy*, 223:124–133.

Hendry, D. F. (2020). First in, first out: Econometric modelling of UK annual CO2 emissions: 1860-2017. *Nuffield College technical report*.

Kaartinen, M. and Prane, O. (2024). Carbon pricing and fuel switching by firms: Theory and evidence. *draft*.

Kaufmann, S. and Strachan, R. W. (2024). Dynamic factor models with common (drifting) stochastic trends. Technical report, Working Paper.

Khan, H., Metaxoglou, K., Knittel, C. R., and Papineau, M. (2019). Carbon emissions and business cycles. *Journal of Macroeconomics*, 60:1–19.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *American Economic Review*, 93(4):1216–1239.

Le Quéré, C., Jackson, R. B., Jones, M. W., Smith, A. J., Abernethy, S., Andrew, R. M., De-Gol, A. J., Willis, D. R., Shan, Y., Canadell, J. G., et al. (2020). Temporary reduction in daily global co2 emissions during the Covid-19 forced confinement. *Nature Climate Change*, 10(7):647–653.

McCulloch, R. E. and Tsay, R. S. (1994). Bayesian analysis of autoregressive time series via the Gibbs sampler. *Journal of Time Series Analysis*, 15(2):235–250.

Miranda-Agrippino, S. and Rey, H. (2020). US monetary policy and the global financial cycle. *The Review of Economic Studies*, 87(6):2754–2776.

Müller, U. K. and Watson, M. W. (2018). Long-run covariability. *Econometrica*, 86(3):775–804.

Mumtaz, H. and Musso, A. (2021). The evolving impact of global, region-specific, and country-specific uncertainty. *Journal of Business & Economic Statistics*, 39(2):466–481.

Mumtaz, H. and Theodoridis, K. (2017). Common and country specific economic uncertainty. *Journal of International Economics*, 105:205–216.

NAEI (2022). NAEI, UK national atmospheric emissions inventory - DEFRA, UK. Available online at `https://naei.beis.gov.uk/`.

Onatski, A. and Wang, C. (2021). Spurious factor analysis. *Econometrica*, 89(2):591–614.

ONS (2022). Developing quarterly greenhouse gas emissions accounts, UK: May 2022. Available online at `https://www.ons.gov.uk/economy/environmentalaccounts/articles/developingquarterlygreenhousegasemissionsaccountsuk/may2022`.

ONS (2023). Environmental accounts on air emissions QMI. Available online at `https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/environmentalaccountsonairemissionsqmi`.

ONS (2024). Review of UK standard industrial classifications 2007. Available online at `https://analysisfunction.civilservice.gov.uk/government-statistical-service-and-statistician-group/user-facing-pages/review-process-for-the-uk-standard-industrial-classification-of-economic-activities-2007-uk-sic-2007`.

ONS (2025). Measuring UK greenhouse gas emissions. Available online at `https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/measuringukgreenhousegasemissions`.

Peña, D. and Poncela, P. (2006). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, 136(4):1237–1257.

Rafaty, R., Dolphin, G., and Pretis, F. (2025). Carbon pricing and the elasticity of CO2 emissions. *Energy Economics*, 144:108298.

Rogelj, J., Geden, O., Cowie, A., and Reisinger, A. (2021). Three ways to improve net-zero emissions targets. *Nature*, 591(7850):365–368.

Saboya, E., Zazzeri, G., Graven, H., Manning, A. J., and Englund Michel, S. (2022). Continuous CH_4 and $\delta^{13}CH_4$ measurements in London demonstrate under-reported natural gas leakage. *Atmospheric Chemistry and Physics*, 22(5):3595–3613.

Stechemesser, A., Koch, N., Mark, E., Dilger, E., Klösel, P., Menicacci, L., Nachtigall, D., Pretis, F., Ritter, N., Schwarz, M., et al. (2024). Climate policies that achieved major emission reductions: Global evidence from two decades. *Science*, 385(6711):884–892.

Van Den Ende, R., Mandel, A., and Rusinowska, A. (2025). Network-based allocation of responsibility for GHG-emissions. *Social Choice and Welfare*, forthcoming.

Wen, H.-x., Chen, Z., Yang, Q., Liu, J.-y., and Nie, P.-y. (2022). Driving forces and mitigating strategies of CO2 emissions in China: A decomposition analysis based on 38 industrial sub-sectors. *Energy*, 245:123262.

Supplementary Information

A Data

In this appendix, we first briefly describe how industry-level emissions data are constructed. This is essential, as this process underpins our theoretical framework. Next, we describe how we process and match the datasets we use.

A.1 Understanding industry-level emissions data

To fulfill their international obligations to report emissions, countries develop National Emissions Inventories, which involves collecting data under process-based frameworks (the SNAP framework or, more recently, the CRF/NFR frameworks).

These nomenclatures have been developed to focus on the process at the origin of the emissions, rather than on the economic industry that conducts this process. As a result, to obtain industry-level emissions data, national accountants need to map these process-based nomenclatures to economic taxonomies such as SIC or NACE – this is known as the "inventory-first" approach, and is used by the ONS and most of Eurostat countries.

This mapping is often not obvious. Eurostat (2009) recommends national accountants to (1) understand exactly how the process-based emissions were calculated; (2) understand exactly how economic industries are defined; (3) make a recommendation for a mapping. In many cases, there is an obvious one-to-one mapping, but in other cases many industries perform the polluting activity. Table 4 provides an example.

Table 4: Procedure for mapping NAEI to SIC07 (ONS, 2023)

Mapping	Allocation Procedure	Examples
One-to-one	A single NAEI activity and source is judged as equivalent to a single SIC 2007 code.	Gas oil for freight trains is matched with the SIC code corresponding to the rail transport industry.
One-to-many	The activity and its emissions are split among the relevant SIC codes either directly (based on the proportionate level of activity) or indirectly (as determined primarily by Ricardo Energy and Environment).	Emissions from motorway driving are allocated proportionally across various SIC industries and households, using data from the National Travel Survey and Purchases data on fuel expenditures.

Eurostat provides a general mapping between the NAEI activities and sources and the industries (NACE Rev.2), but does not offer precise calculations for the allocation process. They indicate, for instance, that the usage of passenger cars for urban driving (SNAP code 07 01 03, CRF/NFR code 1.A.3.b.i) will be assigned to all industries (i.e., 01-99). However, they do not detail the exact fuel sources, quantities, and distribution of miles traveled by cars for urban driving across each relevant industry.

Territoriality. Both NAEI (ONS, 2023) and national emission inventories used by Eurostat (Eurostat, 2009) are on a *territory basis* and follow UNECE Convention on Long-range Transboundary Air Pollution (CLRTAP) with reporting to UNECE/EMEP. For the final emissions data, both Eurostat and ONS are in accordance with the UN System of Environmental Economic Accounts (SEEA), which is on a *residency basis*. The residency principle enables the greenhouse gas emissions (GHG) to be comparable with the National Accounts, including gross domestic product (GDP) (ONS, 2022). Adjustments for the residence principle are typically made first, and then the emissions are allocated to the economic activities.

A.2 Building our datasets

We use three datasets.

(a) Gross value added data is available at:
Main page: `https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/ukgdpolowlevelaggregates/current`
Dataset: `https://www.ons.gov.uk/file?uri=/economy/grossdomesticproductgdp/datasets/ukgdpolowlevelaggregates/current/gdplowlevelaggregates2024q4.xlsx`; We use the sheet "2a" or "CVM Millions". It covers the period 1990–2024 and 106 SIC industries.

(b) Emissions data are available at the page:
Main page: https://www.ons.gov.uk/economy/environmentalaccounts /datasets/ukenvironmentalaccountsatmosphericemissionsgreen housegasemissionsbyeconomicsectorandgasunitedkingdom
Dataset: https://www.ons.gov.uk/file?uri=/economy/environmentalaccounts/datasets/ukenvironment alaccountsatmosphericemissionsgreenhousegasemissionsbyeconomicsectorandgasunitedkingdom/current/ provisionalatmoshpericemissionsghg.xlsx; The greenhouse gases measured include carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydro-fluorocarbons (HFC), perfluorocarbons (PFC), nitrogen trifluoride (NF_3), and sulphur hexafluoride (SF_6). All are expressed in units of kilotonnes of carbon dioxide equivalent. We use the aggregate CO_2 equivalent. The data cover the period 1990–2022 and 131 industries.

(c) Eurostat concordance table is available at:
`https://ec.europa.eu/eurostat/documents/1798247/6191529/Annex1-AEA-Manual-%28SNAP-CRFNFR-NACE2%29-v2+0.xls`.

We first describe how we match ONS emissions and GVA data. Next, we discuss how we match the Eurostat activities data to our matched emissions-GVA data.

A.2.1 Matching emissions and GVA data

Industry-level data from ONS is classified by Standard Industrial Classification (2007), commonly referred to as "SIC07". Both the GHG and GVA data are classified into over 100 categories at the SIC 2-4 digit level, although there are some differences in the level of disaggregation between these two variables. The GVA data (106 categories) is slightly more aggregated than the GHG data (131 categories). Thus, to calculate emissions intensity, we aggregate the granular SIC sectors in GHG dataset with their exact corresponding categories in the GVA dataset based on the sector categories and names defined in the GVA raw excel file specified in the section above. Table 7 provides a detailed breakdown of the industries.

A.2.2 Matching Eurostat to our data

To create blocks of industries based on their common use of polluting activities, we need to match the Eurostat NACE-based data to our SIC-based dataset. First, since we focus on industry-level emissions in our analysis, we exclude all residential activities in the corresponding table. Second, we match the sectors in the ONS dataset with the ones in the Eurostat concordance table. The sectors in the ONS dataset follow the SIC classification at the 2-4 digit level. Because UK SIC 2007 is the national version of NACE Rev. 2 (both systems are identical down to the 4 digit 'class' level) (ONS, 2024), we manually create a cross-walk between 2-digit industries in the Eurostat table and 2-4 digit level industries in the ONS dataset.

Matching the industries in the Eurostat corresponding table with the ONS dataset gives us direct matches between the emission activities (in SNAP 97 and NFR/CRF codes) and the SIC sectors. We then need to decide whether to use SNAP or NFR/CRF. While the UK NAEI website[12] shows activities classified based on the NFR/CRF coding system (Saboya et al., 2022), the NFR/CRF codes in the Eurostat table are not as detailed as the SNAP categories, which would lead to one-to-many relationships between the polluting activity and the SIC07 industry. Thus, we use the SNAP code instead. This procedure gives us a mapping between each SNAP code in the Eurostat concordance table and one or several SIC07 industries. Fig. 5 provides a visual representation of this mapping.

A.2.3 Emission factors

In the main paper, we assume that emissions per unit of activity are constant. In practice, emission factors typically refer to the amount (in a specific unit, e.g., kilotonne) of greenhouse gas emissions produced by one unit (e.g. TJ) of emissions-producing activity, under a specific fuel source. As the definition of activities is per unit of fuel/source used, generally speaking it is unlikely that the emission factor changes significantly over time. For example, while cars may consume less petrol per km now than in the past, their emissions per unit of petrol burned is likely to be more stable. That said, continuing this example, technologies like catalytic converters would affect how

[12] `https://naei.beis.gov.uk/glossary?view=crf`

much methane per unit of petrol burned is released in the atmosphere. Thus, while the stability of emission factors appears a-priori as a good first-order approximation, emission factors can change and the extent to which they change is an empirical question.

To check this more carefully, we downloaded emission factors from UK NAEI emission factor database[13]. The database provides available emission factors of each activity from 1990 to 2023, for each type of greenhouse gas. We define an *activity series* as a unique tuple of Pollutant, Sector, NFR/CRT Group, Source, and Fuel Name. The raw file contains 2271 unique activities.

As an imperfect attempt to keep activities that are most likely to be related to industries which we analyse in our paper, we exclude rows where (1) the Sector is "Land-Use, Land-Use Change and Forestry", as these are likely to mostly be unrelated to industrial activities; (2) rows where the NFR/CRT Group is "Residential stationary fuel use", since we focus on industry (i.e. household emissions are excluded from our analysis); and (3) rows with empty "Activity Units", as we suspect these may represent emissions amounts that are not normalised per unit of activity. This leaves us with 1850 unique activities. For simplicity we also remove activities for which less than 34 years of data are available, which brings us to 1653 unique activities. Finally, we remove activities that have an emission factor of exactly zero throughout the sample, leaving us with 1617 activities as our final, balanced, panel sample. We recognise that these cleaning steps remove emissions factors that are highly volatile, and that the final sample does not match perfectly the boundaries of emissions included in the ONS industry-level emissions, but we think these are reasonable cleaning steps leading to a more insightful analysis.

With this balanced panel, we analyse the year-on-year stability of NAEI activity-level emission factors using annual log growth rates. For each retained activity a, let $\theta_{a,t}$ denote the reported emission factor in year t. We compute log differences[14] for all series. Some series may contain the value 0, so we remove non-finite growth rates. We then compute the absolute value of the mean growth rate ("drift"), and the standard deviation of growth rates ("volatility").

		$= 0$	$(0, 0.01]$	$(0.01, 0.1]$	$(0.1, 0.5]$	$(0.5, 0.8]$
$\lvert\mu_a\rvert$	Count	700	773	109	35	0
	%	43.29	47.80	6.74	2.16	0.00
	Cumul. %	43.29	91.09	97.84	100.00	100.00
σ_a	Count	700	621	210	58	28
	%	43.29	38.40	12.99	3.59	1.73
	Cumul. %	43.29	81.69	94.68	98.27	100.00

Table 5: Distribution of $|\mu_a|$ and σ_a across the 1,617 retained activity series.

Table 5 shows the distribution of these quantities. The distribution is highly concentrated near zero: roughly 90% of series have a drift $|\mu_a| \leq 1\%$ per year, and 80% have a volatility $\sigma_a \leq 1\%$,

[13] `https://naei.energysecurity.gov.uk/emission-factors/emission-factors-database`.

[14] In the main paper we use absolute difference, but here activities' emission factors are in different units, so for comparability we measure changes in relative terms.

indicating that the large majority of NAEI emission factors are effectively constant over time. There is a tail of activities with larger drift and volatility. While a detailed analysis of these volatile emissions factors is outside the scope of this paper, Table 6 shows the top five most volatile series.

Pollutant	Sector	NFR/CRT group	Source	Fuel name	Unit	Activity unit	Mean EF	Drift	Vol.
Methane	Energy	Oil and gas extraction fuel use	Gas terminal: fuel combustion	Gas oil	kT	TJ (net)	5.31e-06	0.01	0.77
Bio-Carbon	Energy	Agriculture/ Forestry/ Fishing Stationary fuel use	Agriculture - stationary combustion	Natural gas	kT	TJ (net)	2.63e-05	0.46	0.67
Bio-Carbon	Energy	Chemicals industries fuel use	Chemicals (combustion)	Natural gas	kT	TJ (net)	2.63e-05	0.46	0.67
Bio-Carbon	Energy	Commercial/ Institutional stationary fuel use	Miscellaneous industrial/commercial combustion	Natural gas	kT	TJ (net)	2.63e-05	0.46	0.67
Bio-Carbon	Energy	Commercial/ Institutional stationary fuel use	Public sector combustion	Natural gas	kT	TJ (net)	2.63e-05	0.46	0.67

Table 6: Top five activities in terms of std. dev. of log growth rates. EF stands for Emission Factor, and Vol. for Volatility (std. dev. of growth rates). Note that several activity labels share identical emission-factor trajectories.

Interestingly, these concern mostly Bio-Carbon pollutants (a list longer than five would also be dominated by Bio-Carbon). The column "Mean Emission Factors" shows that emission factors for these activities are "small", in the sense that the absolute values are of the order $\sim 10^{-5}$ kT/TJ, although it is difficult to assess in a systematic manner whether these activities would matter to larger aggregates without knowing the amount of activity carried out, and the global warming potential of the pollutant. That said, we think many of the highly volatile emission factors are for pollutants that represent a small share of the total CO_2e totals. In the main paper, we work with CO_2e totals, which are likely to be much more stable due to aggregation.

Overall, while it is clear that technological change, regulation and other factors may affect emission factors, this analysis confirms that assuming constant emission factors is likely to be a good first-order approximation.

B Clustering industries into blocks using community detection

We use a classic network community detection algorithm as a way to create non-overlapping blocks such that industries inside a block tend to have a lot of common activities. First, starting from the industries-activities binary bipartite graph, we create the one-mode projection, where nodes are industries and links are the number of activities that they have in common. Next, we use a standard community detection algorithm, the Louvain method (Blondel et al., 2008). We choose it because it is well understood and has a "resolution" parameter, which makes it possible to observe communities at different scales or hierarchies. This allows us to choose a scale corresponding to a rather small number of communities – we were looking for a number between, say, 3 and 10, because the number of blocks in the block-level factor model should be neither too large nor too small.

The Louvain method allows us to decide roughly how many communities to have, while letting the exact number be guided by the data; specifically, it is a number of communities that remains optimal over a wide range of values of the resolution parameter.

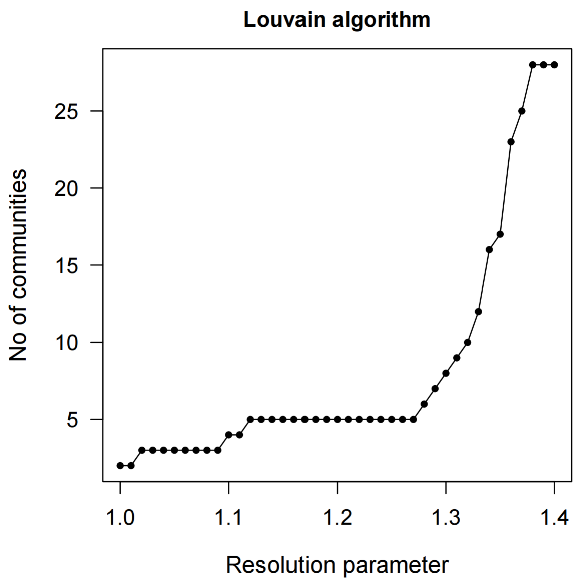

Figure B.1: The number of communities vs. resolution parameter of the Louvain algorithm.

Fig. B.1 shows the number of communities as a function of the resolution parameter. There is a long plateau at 5 communities. We choose the partition found at resolution value 1.2. Table 7 shows the details of the communities (see also Fig. 5 in the main text).

Table 7: Blocks and industries.

Blocks	No. of industries	Industries
S1	3	A 01 Agricultural products
		A 02 Forestry products
		A 03 Fishery products
S2	7	B 05 Coal mining
		B 06 Crude oil & gas extraction
		B 08 Other mining & quarrying
		B 09 Mining support
		C 19 Coke oven & petroleum products
		D 35.1 Electricity production (gas, coal, nuclear, oil, other)
		D 35.2-3 Gas & steam supply
S3	42	C 10.1 Meat production
		C 10.2-3 Fish, seafood, fruit & veg processing
		C 10.4 Oils & fats production
		C 10.5 Dairy products
		C 10.6 Grain milling

		C 10.7 Bakery & farinaceous products
		C 10.8 Other food products
		C 10.9 Animal feeds
		C 11.01-06&12 Alcoholic drinks and tobacco products
		C 11.07 Soft drinks & bottled water
		C 13 Textiles
		C 14 Apparel
		C 15 Leather & related products
		C 16 Wood products (excl. furniture)
		C 17 Paper products
		C 18 Printing & recording
		C 21 Pharmaceuticals
		C 22 Rubber & Plastics products
		C 23.1-4 & 23.7-9 Glass & ceramics
		C 23.5-6 Cement, lime, plaster & concrete products
		C 24.1-3 Basic iron & steel
		C 24.4-5 Aluminium, nuclear fuel, other metals & casting
		C 25 Manufacture of fabricated metal products, except machinery and equipment
		C 26 Computers & electronics
		C 27 Electrical equipment
		C 28 Machinery
		C 29 Motor vehicles
		C 30.1 Ships & boats
		C 30.2+4+9 Other transport (excl. air, space)
		C 30.3 Air & spacecraft
		C 31 Furniture
		C 32 Other manufacturing
		C 33 (not 33.15-16) Repair & installation (excl. ships & aircraft)
		C 33.15 Ships repair
		C 33.16 Aircraft & spacecraft repair
		E 36 Water supply
		E 37 Sewerage
		E 38 Waste management
		E 39 Remediation & other waste services
		F 41 Buildings & construction
		F 42 Civil engineering
		F 43 Specialised construction
S4	6	C 20A Industrial gases, fertilisers & chemicals
		C 20B Petrochemicals
		C 20C Dyes & agrochemicals
		C 20.3 Paints & inks
		C 20.4 Cleaning & toiletries
		C 20.5 Other chemicals
S5	48	G 45 Motor vehicle trade & repair
		G 46 Wholesale trade (excl. motor vehicles)
		G 47 Retail trade (excl. motor vehicles)
		H 49.1-2 Rail transport
		H 49.3-5 Land transport
		H 50 Water transport
		H 51 Air transport
		H 52 Warehousing & support services
		H 53 Postal & courier services
		I 55 Accommodation

I 56 Food & drink services
J 58 Publishing
J 59 Film, video & music production
J 60 Broadcasting
J 61 Telecommunications
J 62 IT services
J 63 Information services
K 64 Financial services (excl. insurance)
K 65.1-2 Insurance & reinsurance
K 66 Auxiliary financial services
L 68.1-2 Real estate (excl. imputed rent)
L 68.3 Real estate on a contract basis
M 69.1 Legal services
M 69.2 Accounting & tax consultancy
M 70 Head offices & management consultancy
M 71 Architecture & engineering
M 72 R&D
M 73 Advertising & marketing
M 74 Photography, translation & other activities
M 75 Veterinary services
N 77 Leasing
N 78 Employment agencies
N 79 Travel agents & tour operators
N 80 Security
N 81 Buildings & landscape care
N 82 Office support
O 84 Public administration
O 85 Education services
Q 86 Human health services
Q 87 Residential care services
Q 88 Social work services without accommodation
R 90 Creative & arts
R 91 Library & other cultural services
R 92 Gambling & betting services
R 93 Sporting services
S 94 Services furnished by membership organisations
S 95 Repair services of personal and household goods
S 96 Other personal services

The first community is 01–03, which are all agriculture-related industries. The second block includes industries 05–09, 19 and 35, which are related to oil, gas, electricity and petroleum production. The third block includes industries 10.1–43 which are general construction, waste and manufacturing industries such as electric products except the chemical-production related industries. The fourth block includes all the chemical-related manufacturing industries, such as fertilisers and agro-chemicals. Finally, the fifth block includes all the transportation and service industries from industry 45 to 96.

This classification is intuitive, yet non-trivial as it is different from splitting according to upper layers of the ISIC system, see Fig. 5.

The Louvain algorithm is stochastic, but different random seeds give very similar results. We have also checked another algorithm ("Leiden"), and found broadly similar results.

C Comparison of the block-level vs. global-only factor model

In this section, we compare the results of our block-level model with those of a standard factor model.

C.1 The standard approximate factor model

$$X_{it} = l_i f_t + \nu_{it} \tag{26}$$

where X_{it} is the standardised first-difference of emissions intensity (Eq. 15 in the main text), l_i are the loadings, f_t is the common factor, and ν_{it} is the idiosyncratic shock. Every year, a scalar f_t impacts all the industries, although to a different extent, determined by their loading l_i. All variation that is not the result of this is due to a remaining term, which is idiosyncratic (indexed by i and t).

We estimate the basic factor model using the PCA method, specifically as described in Bai (2003). Denote the $N \times T$ data matrix X and rewrite Eq. 26 in matrix form as

$$X = \Lambda F' + V,$$

where Λ is $N \times 1$, F is $T \times 1$, and V is $N \times T$. This formulation makes it clear that the factor model can be seen as decomposing the panel data matrix as a matrix of noise plus a rank-1 perturbation – a "spiked" random matrix.

The common factor is estimated as $\sqrt{T}$ times the main eigenvector of the $T \times T$ matrix $X'X$, and denoted $\tilde{F}$. Once we have the factors, the loadings are estimated by regressing the data on the factors, that is $\tilde{\Lambda} = (\tilde{F}'\tilde{F})^{-1}\tilde{F}'X' = \tilde{F}'X'/T$ where the last step comes from the assumption that $F'F/T = I$, that is, $\mathrm{Var}(f_t) = 1$. Bai (2003) developed inferential theory for these estimators under general assumptions, and in particular allowing for cross-sectional dependence of the idiosyncratic terms, so that the model is an "approximate" factor model.

From Eq. 26, the variance decomposition is

$$\mathrm{Var}(X_{it}) = l_i^2 \mathrm{Var}(f_t) + \mathrm{Var}(\nu_{it}) = l_i^2 + \mathrm{Var}(\nu_{it}). \tag{27}$$

Thus, the contribution of the common factor to the variance of an industry (and thus to aggregate volatility of emissions intensity) is simply the squared loadings, as in Eq. 20 for the block factor model.

C.2 Conceptual differences between global-only and block factor model

Both models feature a global factor and idiosyncratic noise. The global factor captures a latent factor that is the same for all units, even though industries are heterogeneous in how much of their variance depends on the global factor (through their loadings, see Eq. 27).

In our context, we expect that a global factor would capture all sources of emissions reduction that are linked to macro events, macro policies, or universal inputs (polluting activities undertaken in all industries). A factor model is suitable, because it allows all the industries to be affected, but recognises that they would be affected to a different extent.

However, in the global-only model, the global factor captures all common variation. It is conceivable that some variation is common only to a specific block of industries, but estimating a global-only factor model would attribute this block-specific variation to the global factor (and hence, "macro" policies). Indeed, we have found visually (Fig. 4) that the group of service industries tend to co-move strongly, so there is a concern that a global-only factor model would attribute block-specific variation to the common factor. We have also seen (Fig. 5) that some sets of inputs are common to some specific sets of industries. This suggests using a model that estimates block-specific factors, on which industries outside the block have zero loadings by assumption.

In a block-level model, block-specific factors explain within-block variation. If the block-level model fits better, in the sense that there is indeed variation that is block-specific, the loadings on the global factor should become lower. This would be a sign that the loadings in the global-only model overestimate common variation by ignoring block-specific factors.

In other words, in the block-level model, the global factor coexists with additional block-specific factors. This means that the global factor now captures shared variation across all blocks, while the block-level factors explain the variation within each block. The loadings on the global factor in this model will reflect how much of the total variation each variable still shares with the global trend, after accounting for block-specific variations.

Similarly, the block-level model may explain more variation than the global factor alone, in the sense of having a lower variation explained by the idiosyncratic shocks. This would mean that the block-level factor model provides a better description of the data, in terms of separating what is due to truly global factors, and what is due to more sectoral policies, that may manifest themselves at the detailed industry level (idiosyncratic shocks) or larger industry groups (block-level)

In the next section, we compare the results of the global-only and block-level factor model.

C.3 Comparing the global factors

Fig. C.1 shows the global factor estimated via the global-only model and the block-level dynamic factor model. The two models produce very similar global factors. Thus, the differences between the two models, in terms of the impact of the global factor, will come mostly from the size of the loadings.

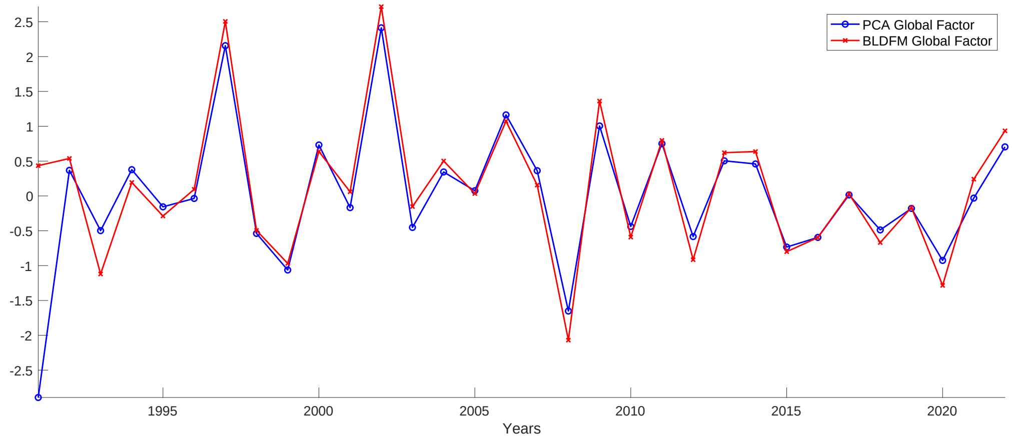

Figure C.1: Comparison of common factor estimated using Principal Component Analysis (PCA) and the proposed Block-level Dynamic Factor Model (BLDFM).

Fig. C.2 shows the global and the block-level factor estimates (cumulative) with the Bayesian credible set. This makes it clear that within-group variation is more significant for some groups than for others.

C.4 Comparing the loadings on the global factors

The bar plot in Fig. C.3 displays the loadings on the global factor from both models, with one bar for the global-only model and another for the block-level model. This provides a systematic way to visually inspect differences across industries.

For ease of reference, it is helpful to remind from Table 7 that the blocks roughly correspond to Agriculture (S1), Energy (S2), Manufacturing (S3), Chemistry (S4) and Services (S5).

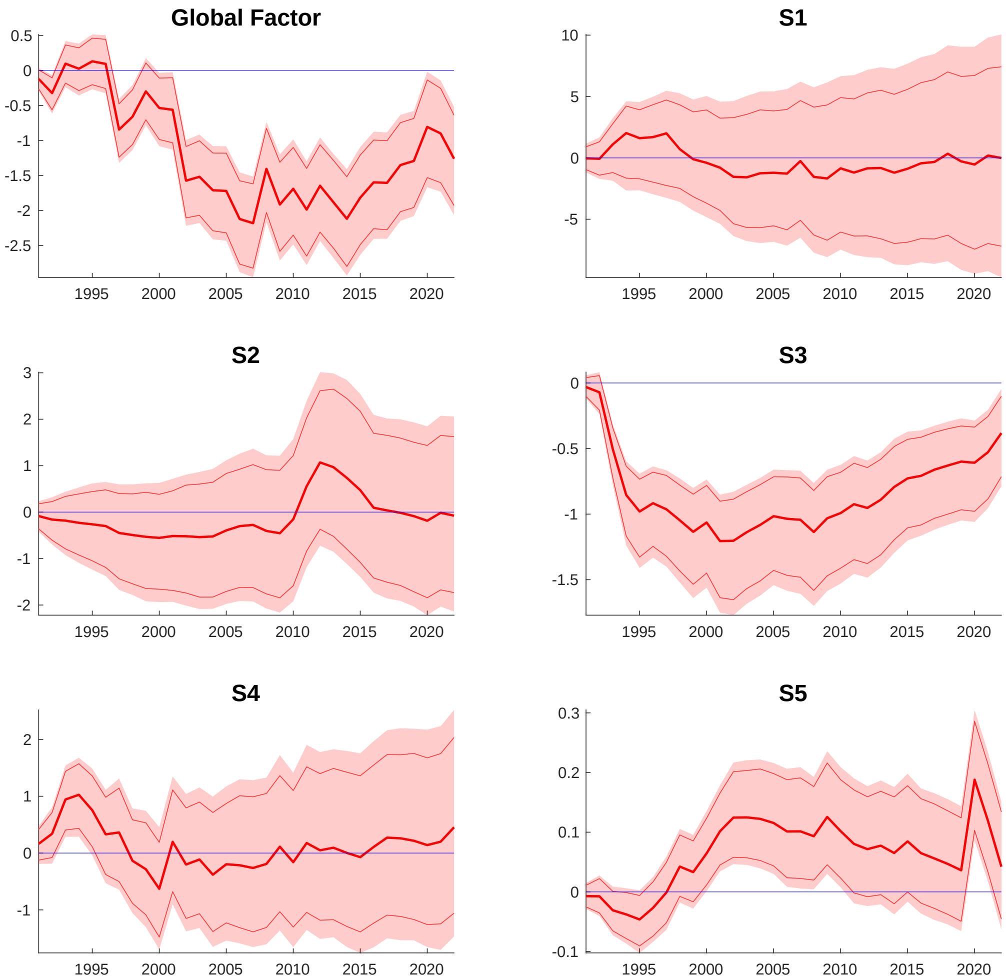

Figure C.2: Time series of the Bayesian estimate of the cumulative factors, along with the 5th and 95th percentiles (shaded region), and the 2.5th and 97.5th percentiles (outer shaded region) of the posterior distribution. The solid line represents the median (50th percentile) estimate of the factors at each time point. The credible intervals provide a measure of dispersion of the posterior distribution around the estimate.

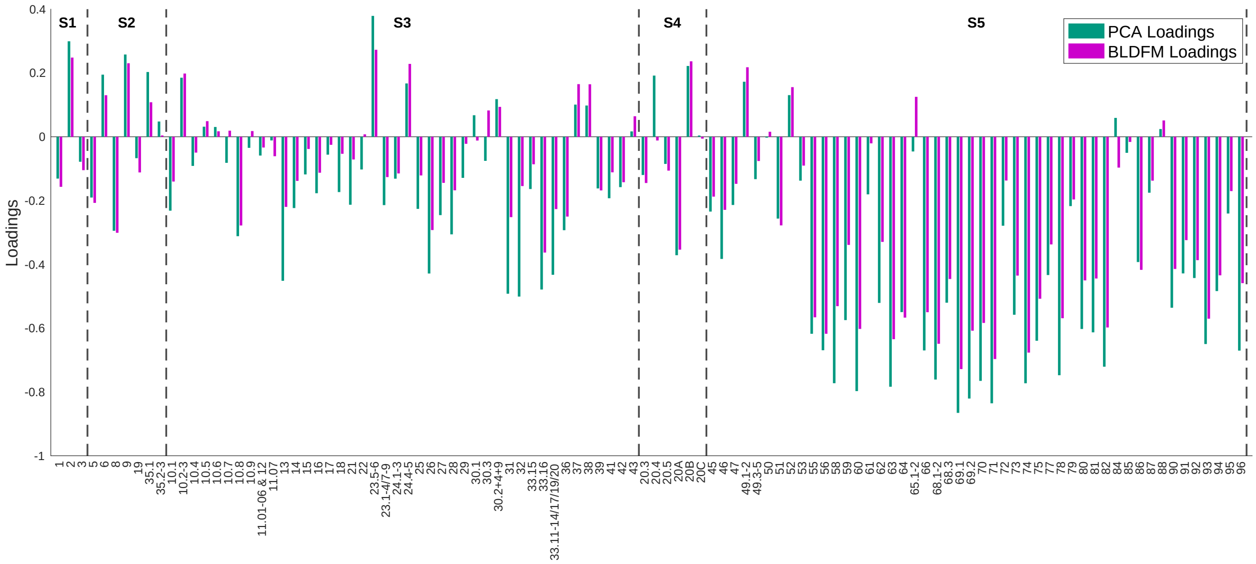

Figure C.3: Comparison of loadings for the common factor estimated using Principal Component Analysis (PCA) and the block-level factor model. The bar chart displays loadings on the common factor for each sector, with PCA loadings shown in teal and BLDFM loadings shown in magenta. Vertical dashed lines separate different blocks.

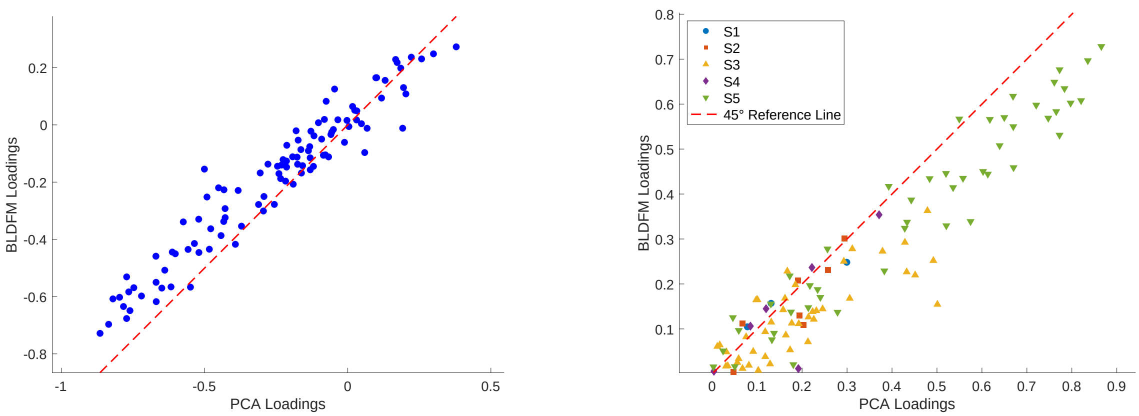

Figure C.4: Comparison of loadings on the common factor from the PCA model and the BLDFM. The left plot displays a scatter plot comparing the PCA loadings with the BLDFM loadings, with a 45-degree reference line indicating where loadings are equal in both models. The right plot shows the same comparison but the loadings are in absolute values and the colors indicate industries from each block assignment, illustrating how loadings differ by block.

A key question is whether the loadings are lower in the case of the block-level model. To check this more easily, the scatter plots in Fig. C.4 compare the loadings from the global-only model (x-axis) with the loadings from the block-level model (y-axis). The red dashed line at 45° serves as a reference — points close to this line indicate similar loadings in both models, while points far from the line show divergence.

On the right panel, we show the absolute values, so that the points below the 45° line indicate that the loadings from the block-level model are smaller than those from the PCA-only model. This suggests that the global factor in the block-level model is explaining less of the variation for these industries than the PCA-only model does. As expected the block-level model assigns less weight to the global factor for Block 5 than the PCA model does, implying that the PCA-only model inflates the importance of the global factor for this block.

C.5 Loadings on the block factors

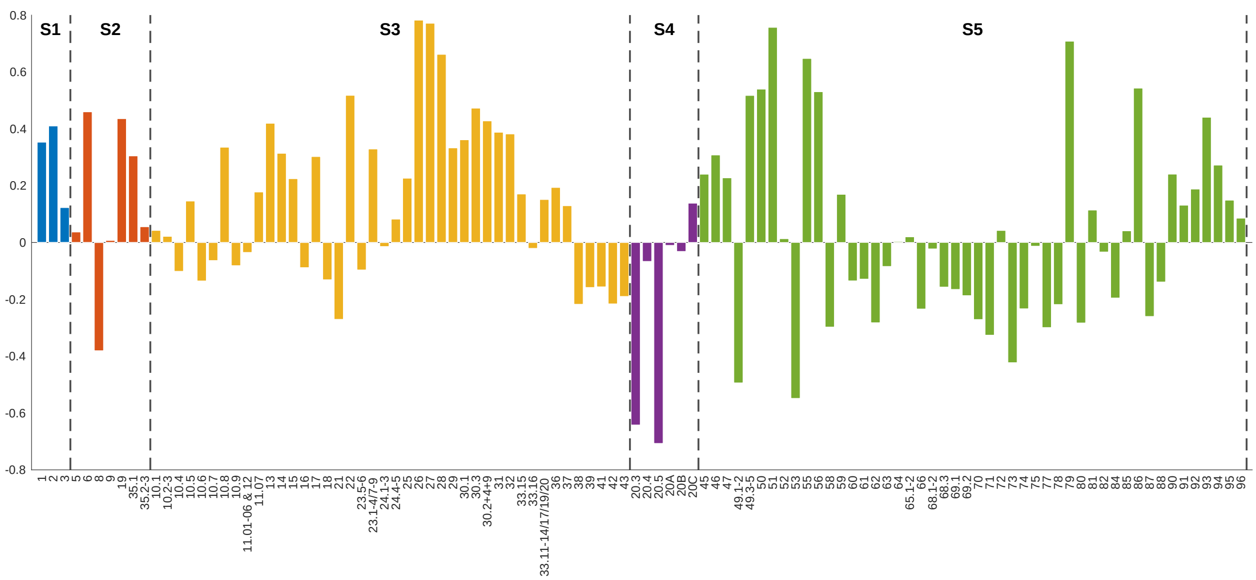

Figure C.5: Loadings of the block-level factors by block. This figure illustrates the rescaled loadings for each block-level factor, with different blocks represented along the x-axis. The vertical bars show the magnitude of the loadings.

Fig. C.5 shows loadings of each block-level factor, with different blocks represented by a different color.

The fact that the service industry (block S5) also has somewhat high loadings on its block-specific factor means that in addition to being influenced by global trends, there are important service-sector-specific dynamics that are not fully captured by the global factor. Similar results hold for the other blocks.

In the block-level model, industries (specifically for the ones in Block S3 and S5) that are better explained by block factors have lower loadings on the global factor, because the global factor no longer needs to explain all the variation. This shows that the block-level factors are capturing block-specific dynamics, leading to a more accurate fit.

The most important case-in-point is S4, which corresponds to chemical manufacturing industries. This community has a very strong block of common activities (Fig. 5), so it is interesting to see that while not all industries have high loadings, all industries have loadings of the same sign, and loadings can be very substantial in some cases (close to 1/2).

C.6 Comparing variance explained

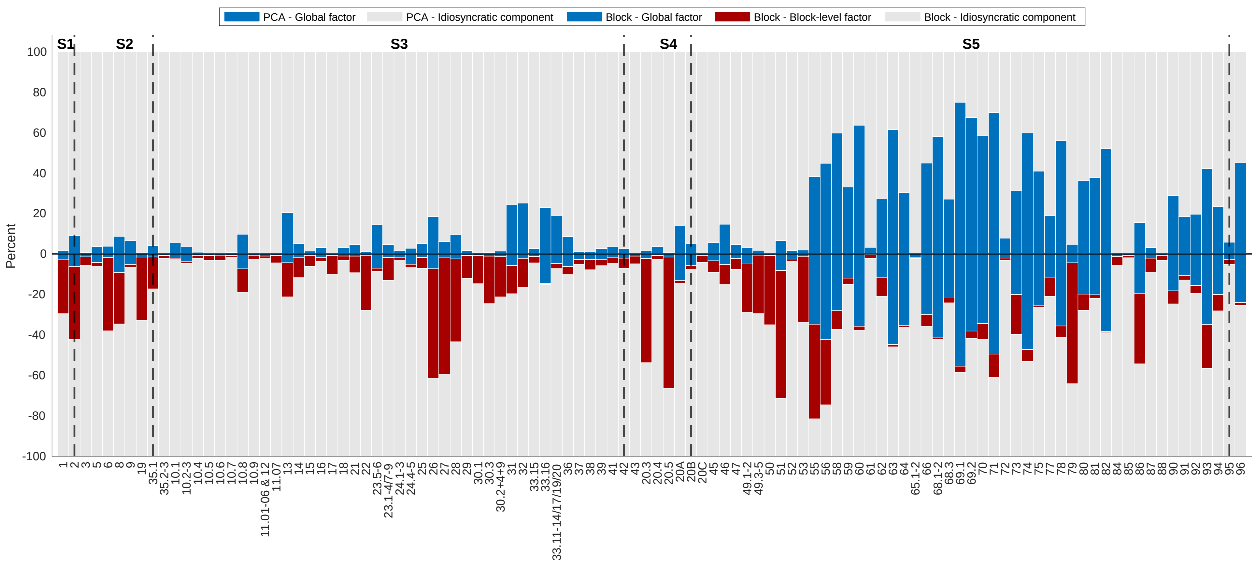

Figure C.6: Comparison of the variance explained by the global factor model with PCA-only (top) and the block-level factor model (bottom). The top part of the figure shows the proportion of variance explained by the principal component analysis (PCA) model across industries, assuming a single global factor. The bottom part of the figure illustrates the variance explained by the block-level factor model, which incorporates additional block-specific factors. This comparison highlights the improved explanatory power of the block-level model in capturing the underlying sectoral structure.

Fig. C.6 compares the variance explained by the two models, showing that the block-level model tends to have a higher explanatory power. To sum this up into one number, the block-level factor model has a lower average residual variance (0.785) compared to the PCA model (0.826).

Fig. C.7 compares the residual variances by block, computed as (unweighted) mean of the residual variances of each industry within the block. It shows that the block-level factor model offers substantial improvements over the PCA model particularly in blocks S1 to S4. Specifically, the residual variances are reduced by approximately 14.5%, 7.4%, 7.5%, and 23.7% in these blocks, respectively. This suggests that the block-level model better captures sector-specific variation and reduces unexplained residuals more effectively than the PCA model. It is interesting to note that the block for which the residual variance is reduced the most is S4 (Chemical Manufacturing), a block with a lot of block-specific activities in which all the industries of the block are involved (Fig. 5)

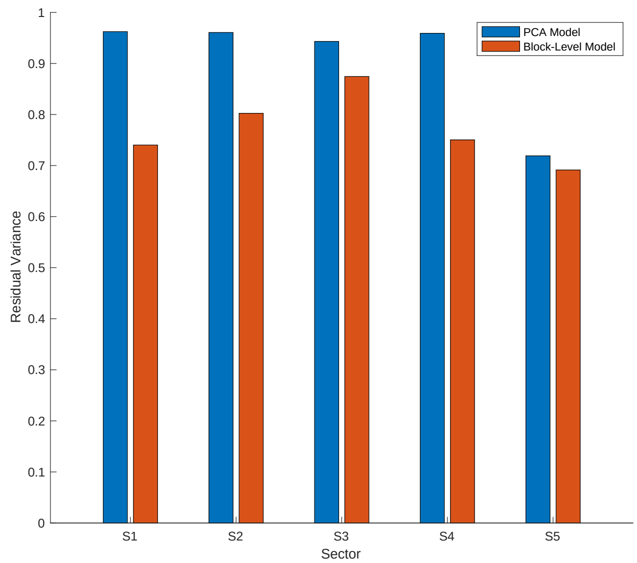

Figure C.7: Comparison of the total residual variance explained by the global factor model with PCA-only (blue bar) and the block-level factor model (orange bar). The figure shows the additional residual variance that the block-level factor model explains compared to the principal component analysis (PCA) model across sectors. This comparison highlights the improved explanatory power of the block-level model in reducing the variance attributed to the idiosyncratic component.

C.7 Contributions to the within effect

In Section 4.2, we compute the contribution of the global factor to the variance of the within effect. Here we repeat the exercise for the simpler global-only factor model, see Fig. C.8, to be compared with Fig. 7

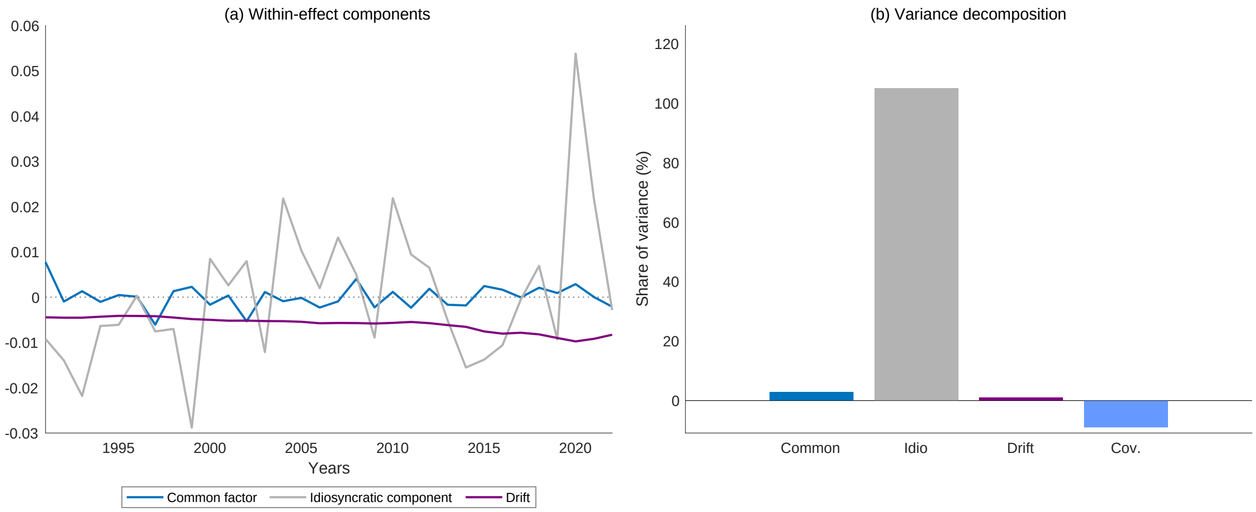

Figure C.8: Analysis of contributions to the within effect. The first subplot illustrates the time series of contributions from the common factor, idiosyncratic component, and drift to the within effect of the emissions data. These contributions are estimated with the PCA-only model. The second subplot presents a stacked bar chart of the variance contributions.

Comparing Fig. C.8 to the equivalent for the block-level model[15], Fig. 7, we find the global factor is very small in both cases. This shows that allowing for block-level common variation, which takes place in key high-emissions sets of industries, is essential to explain aggregate variation. The fact that the common factor contributes so little to the aggregate, despite being significant, also makes clear that looking at loadings of individual industries can be misleading – it is important to re-aggregate the individual contributions properly, by de-standardizing and using industry sizes as weights.

D Interpreting factor models in light of the activities-to-industries emission allocation process

In this section, our primary objective is to provide an interpretation of the common factors. To do this, we establish a connection between the allocation-of-activities DGP and factor models under specific assumptions.

We begin by showing what assumptions are needed to equate the dynamics of emissions intensity with a factor model, showing that in the general case, a factor model does not map exactly with the activities-based data generating process introduced in the main text. We then present special cases where the activities-based DGP can indeed be written as a factor model. We conclude by discussing

[15]The differences between the two models are not only due to the use of blocks, but also in (likely small) part to the fact that the block-level model is Bayesian and allows for autocorrelation.

how this exercise allows us to interpret the co-movement of emission intensities across industries as reflecting both shared activities among industries, but also co-movement between activities.

D.1 The model after demeaning and scaling

We first address the standardization issue, where Eq. 10 describes the dynamics for the change in emissions intensity, but factor models operate on the demeaned and, generally and in our case, on the rescaled data

$$X_{it} = \frac{\Delta\mathcal{E}_{it} - \mu_i}{\sigma_i}, \tag{28}$$

where μ_i and σ_i are the sample mean and standard deviation of the time series $\Delta\mathcal{E}_{i1}, \Delta\mathcal{E}_{i2}, \ldots \Delta\mathcal{E}_{iT}$. We will treat these two moments as fixed constants, acknowledging this is a limitation of the derivation.

Let us start by rewriting the two shocks (u_{jt} for the improvement in industry-specific activity requirement due to activities, and ϵ_{ijt} for the improvement in industry-specific activity requirement due to the industry) as their mean plus the deviation around the mean,

$$u_{jt} = \bar{u}_j + \tilde{u}_{jt}, \qquad \text{Mean}(\tilde{u}_{jt}) = 0,$$

$$\epsilon_{ijt} = \bar{\epsilon}_{ij} + \tilde{\epsilon}_{ijt}, \qquad \text{Mean}(\tilde{\epsilon}_{ijt}) = 0.$$

Using these in Eq. 9, we find that the sample mean of $\Delta\mathcal{E}_{it}$ is

$$\mu_i = \sum_j \theta_j \beta_{ij} \bar{u}_j + \sum_j \theta_j \bar{\epsilon}_{ij}. \tag{29}$$

and therefore

$$\Delta\mathcal{E}_{it} - \mu_i = \sum_j \theta_j \beta_{ij} \tilde{u}_{jt} + \sum_j \theta_j \tilde{\epsilon}_{ijt}. \tag{30}$$

We assume that the σ_is are known and independent numbers. We then have

$$X_{it} = \sum_j \theta_j \beta^*_{ij} \tilde{u}_{jt} + \sum_j \theta_j \tilde{\epsilon}^*_{ijt} = \sum_j \theta_j \beta^*_{ij} \tilde{u}_{jt} + w^*_{it}, \tag{31}$$

where $\beta^*_{ij} \equiv \beta_{ij}/\sigma_i$ and $\tilde{\epsilon}^*_{ijt} \equiv \tilde{\epsilon}_{ijt}/\sigma_i$, which is of the form required. We need to prove two things:

$$\text{Cov}(w^*_{it}, w^*_{i't}) = 0, \quad \forall i \neq i', \forall t, \tag{32}$$

and

$$\text{Cov}(\tilde{\epsilon}^*_{ijt}, \tilde{u}_{jt}) = 0, \quad \forall i, \forall j. \tag{33}$$

As $w^*_{it} = \sum_j \theta_j \tilde{\epsilon}^*_{ijt}$, $\text{Cov}(w^*_{it}, w^*_{i't})$ can be written as $\text{Cov}(\sum_j \theta_j \tilde{\epsilon}^*_{i'jt}, \sum_j \theta_j \tilde{\epsilon}^*_{ijt})$. Based on Eq. 8 and the assumption that σ_i are taken as known and independent numbers, Eq. 32 holds. Meanwhile, Eq. 33 can be directly derived from assumption Eq. 7 and the fact that σ_i as known and independent numbers. Thus, we can avoid notational clutter, and in the rest of the section, we keep working with model (9).

D.2 General case

To recall, our model is (main text Eq. 10) $\Delta\mathcal{E}_{it} = \sum_j \theta_j \beta_{ij} u_{jt} + \omega_{it}$ with the assumption that the two terms are orthogonal. Since a dynamic factor model with K common factors is written

$$\Delta\mathcal{E}_{it} = \sum_{k=1}^{K} B_{ki} F_{kt} + \nu_{it}, \tag{34}$$

if we assume that the idiosyncratic components are the same in the two models ($\nu_{it} = \omega_{it}$), then we have

$$\sum_{k=1}^{K} B_{ki} F_{kt} = \sum_j \theta_j \beta_{ij} u_{jt}. \tag{35}$$

Note that the factors do not necessarily correspond to activities, so the number of factors (K) may well be different from the number of activities. The key assumption is that all the comovement induced by activities can be captured by a parsimonious set of K factors (which need not correspond one-to-one to activities). In general, this would be an untenable assumption. But as we show next, it holds in special cases, illuminating what common factors likely capture in industry-level emissions intensities data.

D.3 Special cases

Single activity Consider the case of a single activity in the economy, e.g. tons of diesel used for car travelling. In this case, we can rewrite Eq. 10 as

$$\Delta\mathcal{E}_{it} = \theta\beta_i u_t + w_{it}, \tag{36}$$

where θ is the emission factor of burning diesel, u_t is the common shock on the tons of diesel used by cars in the economy, β_i is how the common shock u_t affects the activity requirement of industry i. We can rewrite $u_t = F_t$, $\theta\beta_i = B_i$ and $w_{it} = \nu_{it}$. As we have $\mathrm{Cov}(w_{it}, w_{st}) = \mathrm{Cov}(w_{it}, u_t) = 0, \quad \forall i \neq s$, we have $\mathrm{Cov}(\nu_{it}, \nu_{st}) = \mathrm{Cov}(\nu_{it}, F_t) = 0$, which are properties of a factor model. Thus,

$$\Delta\mathcal{E}_{it} = B_i F_t + \nu_{it}. \tag{37}$$

In this case, the common factor F_t is the common shock u_t, and the loading B_i is equal to the emission factor of diesel times the extent to which the GVA of sector i will be affected by this common shock u_t. $\nu_{i,t}$ models industry specific actions in reducing its emission intensity, which is unrelated to this common shock.

The industry-level impact of common shocks to activity requirements is the same across all activities The coefficient β_{ij} measures the extent to which industry i reacts to common shocks in activity requirements u_{jt} (these shocks are common in the sense that they are the same for all industries). If industry i's reaction is the same for all activities j, such that $\beta_{ij} = \beta_i \ \forall j$, we recover again a factor model. Indeed, in this case, Eq.10 can be written as

$$\Delta\mathcal{E}_{it} = \beta_i \sum_j \theta_j u_{jt} + w_{it} = \beta_i F_t + w_{it}, \tag{38}$$

since $\sum_j \theta_j u_{jt}$ only varies across time and can thus be viewed as a common factor F_t.

D.3.1 Interpretation of the co-movement

To better understand the nature of the common factors that would be estimated, we start by calculating the covariance (interpreted as co-movement) of the first differences of emission intensities across industries. Under the assumption that factors F_{kt} are orthogonal, we have

$$\text{Cov}(\Delta\mathcal{E}_{it}, \Delta\mathcal{E}_{st}) = \sum_{k=1}^{K} B_{ki} B_{ks} \text{Var}(F_{kt}), \quad \forall i \neq s. \tag{39}$$

This can be compared to what we get in the activities-based DGP, where the covariance between industry i and s can be expressed as

$$\text{Cov}(\Delta\mathcal{E}_{it}, \Delta\mathcal{E}_{st}) = \text{Cov}\big(\sum_j \theta_j \beta_{ij} u_{jt} + \nu_{it},\ \sum_j \theta_j \beta_{sj} u_{jt} + \nu_{st}\big), \quad \forall i \neq s. \tag{40}$$

Thanks to assumption (5) and (11) we can simplify as follows

$$\text{Cov}(\Delta\mathcal{E}_{it}, \Delta\mathcal{E}_{st}) = \underbrace{\sum_{j \neq m} \theta_j \theta_m \beta_{ij} \beta_{sm} \text{Cov}(u_{jt}, u_{mt})}_{\text{contrib. of cov. between activity requirements}} + \underbrace{\sum_j \theta_j^2 \beta_{ij} \beta_{sj} \text{Var}(u_{jt})}_{\text{contrib. of common activities}}, \ \forall i \neq s. \tag{41}$$

where the overall covariance can be expressed as the sum of two different components: the co-movement among activities, and the co-movement resulting from the common activity allocation. The co-movement among activities refers to the covariance arising from $\text{Cov}(u_{jt}, u_{mt})$, which is the covariance between two different activities (j and m) at time t; and the co-movement resulting from the common activity allocation refers to the covariance originating from $\text{Var}(u_{jt})$, which is the variance of the shock on activity j at time t, but shared between sector i and s.

Then, according to Eq. 41, we can express the right hand side of Eq. 39 as

$$\sum_{k=1}^{K} B_{ki} B_{ks} \text{Var}(F_{kt}) = \underbrace{\sum_{j \neq m} \theta_j \theta_m \beta_{ij} \beta_{sm} \text{Cov}(u_{jt}, u_{mt})}_{\text{covariance among activities}} + \underbrace{\sum_j \theta_j^2 \beta_{ij} \beta_{sj} \text{Var}(u_{jt})}_{\text{variance of common activity allocation}}, \quad \forall i \neq s. \tag{42}$$

Eq. 42 represents the main results of this section. It reveals that if a factor analysis can be utilized to identify the industry-specific idiosyncratic components ν_{it}, which are the emission

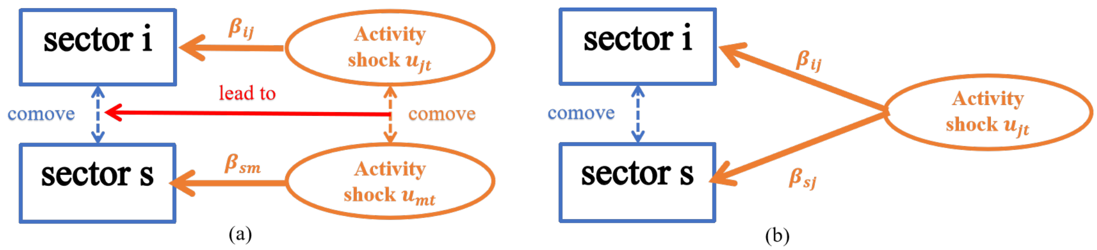

Figure D.1: Visual representation of two different components of the overall covariance

reductions efforts by construction uncorrelated with other industries, the common factors (F_{kt}) will incorporate both information on the shock of the shared activity (u_{jt}), which we show as (b) in Fig. D.1, and the co-movements among activities ($\text{Cov}(u_{jt}, u_{mt})$), which we show as (a) in Fig. D.1. Even when industries share no activities (i.e., $\beta_{ij}\beta_{sj} = 0 \ \forall j$), common factors can still exist in the economy (as $\sum_{j \neq m} \theta_j \theta_m \beta_{ij} \beta_{sm} \text{Cov}(u_{jt}, u_{mt})$ can be non-zero).

The results highlight the complex nature of common activities within the common factors. They indicate that the emission reduction effect from the shocks on the common activities may be shaped by the co-movement among the activities. Meanwhile, the shocks on the idiosyncratic components can have a direct contribution to the industrial emission reduction outcome.

E Correlation matrices by sub-period and COVID-19 robustness check

In the main text, we show that the correlation matrix features block-level co-movement in Section 2.4. To investigate whether synchronization in emissions intensity has changed over time, we split the sample into two subperiods and re-estimate the model separately. This approach also allows us to assess the robustness of our results to the exclusion of the COVID-19 period, as the second subsample ends in 2019.

Figure E.1 presents the correlation matrices of emissions and emissions intensity across subperiods. When considering emissions, there is some evidence of stronger co-movement in carbon-intensive industries. However, for emissions intensity, no clear pattern of increasing synchronization emerges. Figure E.2 provides a quantitative decomposition, showing that the variance explained by both global and block-level factors does not change substantially across subperiods, with the exception of a moderate increase in the importance of the block-level factor for the services sector. Figure E.3 complements this analysis by focusing on the variance decomposition across the two subperiods, both of which exclude the COVID-19 years. The results show that the contribution of the global factor remains broadly unchanged, while the block-level factor exhibits a modest increase for the services block in the later period.

Taken together, these findings suggest that the patterns of co-movement we identify are not driven by pandemic-specific dynamics. Rather, the COVID-19 shock appears to amplify pre-existing

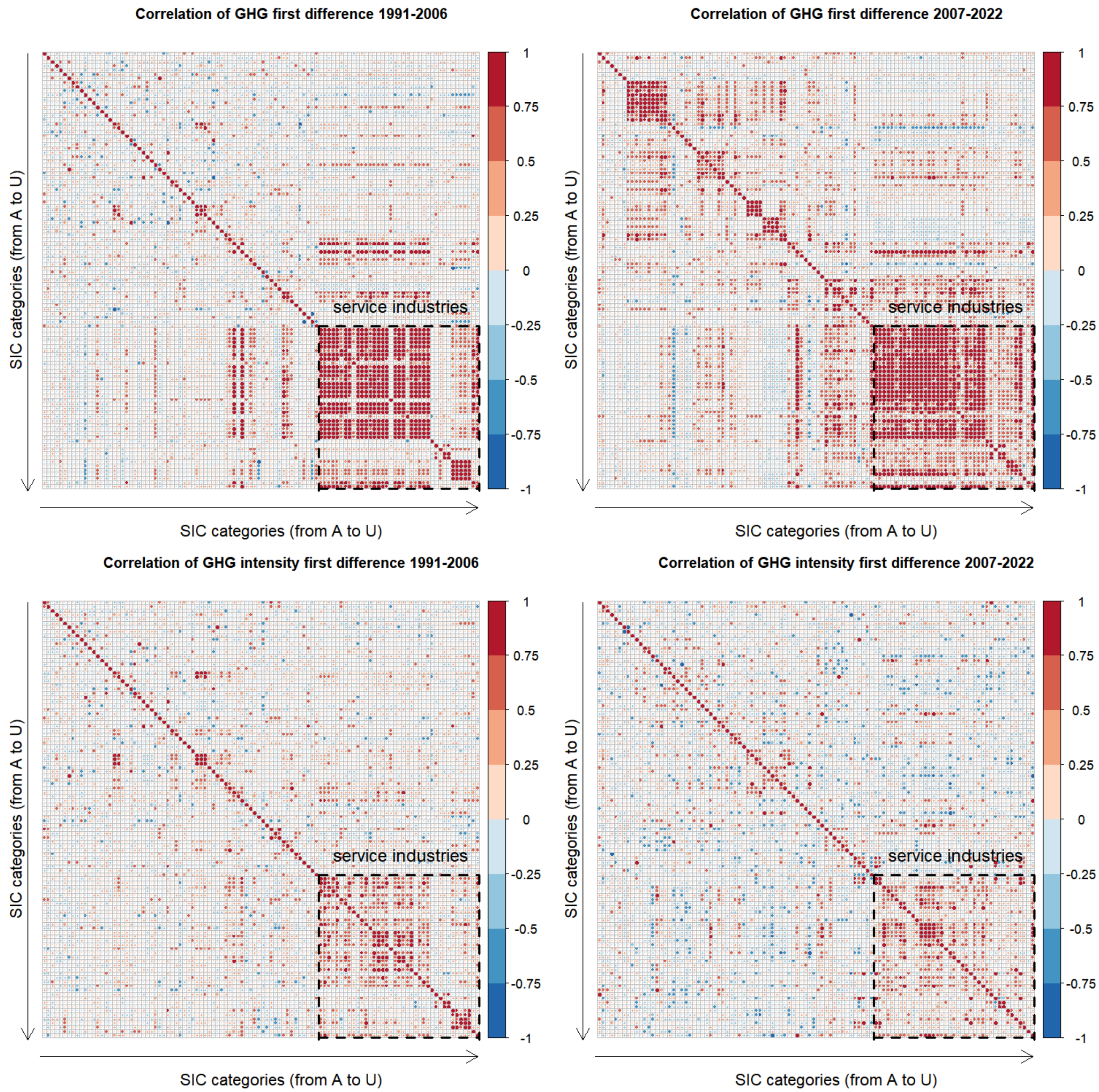

Figure E.1: Correlation matrices for the first differences of emissions and emissions intensity.

patterns of synchronization, without fundamentally altering the underlying factor structure.

Given the relatively short time dimension of the data, detecting changes in synchronization remains inherently challenging, and these results should be interpreted with caution.

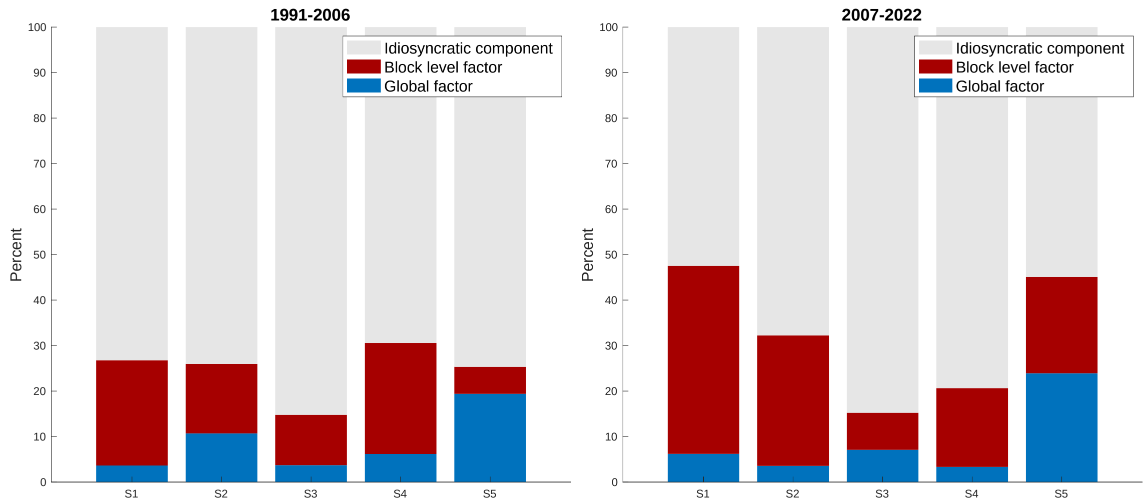

Figure E.2: Proportion of variance explained by the global and block-level factors for two distinct periods: 1991-2006 and 2007-2022. The results are presented at the block level for better visibility. The proportions are calculated individually for each industry and then averaged across the industries within each block.

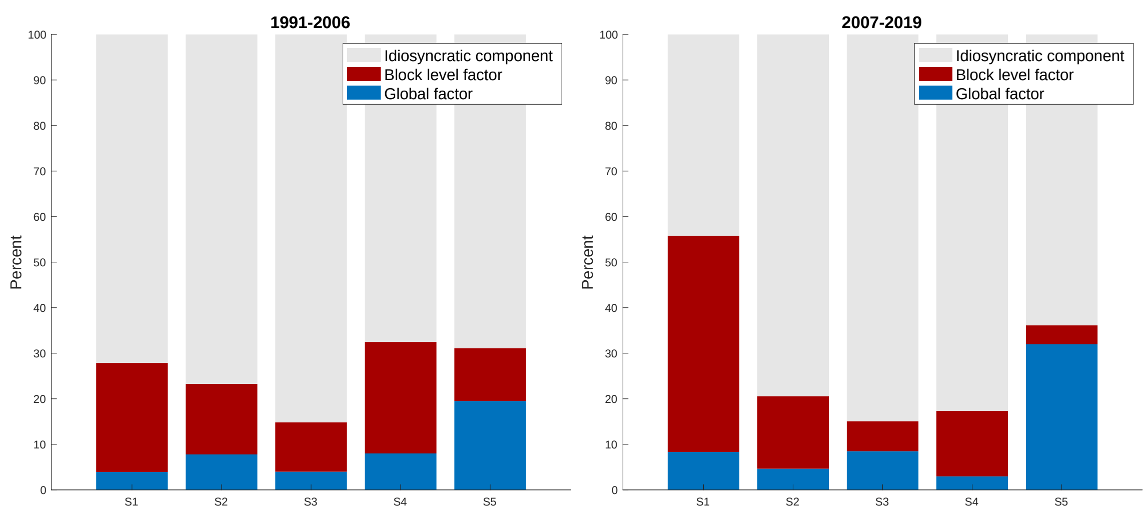

Figure E.3: Proportion of variance explained by the global and block-level factors for two distinct periods: 1991-2006 and 2007-2019, excluding COVID-19 period. The results are presented at the block level for better visibility. The proportions are calculated individually for each industry and then averaged across the industries within each block.