Carbon Emissions in the US:

Factor Decomposition and Cross-State Inequality Dynamics

Panos K Pouliasis*, Nikos C Papapostolou, Michael N Tamvakis and Ioannis C Moutzouris

Bayes Business School, City, University of London

Abstract

This paper examines the determinants of inequality in the distribution of CO_2 emissions across US regions. We implement a factorial decomposition of CO_2 per capita based on extended Kaya factors, that is, carbon intensity of fossil fuel consumption, energy mix, energy intensity of GDP, economic growth in terms of labor productivity and employment rate. Results reveal that US states display marked differences in most factors. We identify energy intensity as the main source of emissions inequality. Based on the within and between group inequality components we also explore the effect of geographical, geological, climatic and human development partitions of US states' groups. Findings indicate that the within-group inequality had been the main contributor to the whole inequality. Finally, some economic policy implications are also discussed; explaining the unequal distribution of emissions is vital to establish differentiated targets and work towards successful mitigation proposals.

Keywords: CO_2 emissions, Theil Index, inequality decomposition, Kaya factors

JEL Classification: Q59, F00, C10

* Corresponding author: Panos K Pouliasis, Bayes Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ. E-mail: panos.pouliasis@city.ac.uk.

1. Introduction

In recent years, a number of international initiatives and abatement methods have been postulated to condense emissions from anthropogenic sources, such as the 1997 Kyoto Protocol and the 2015 Paris Agreement. Energy use from fossil fuels, population growth, technological progress and flexibility to switch from fossil fuels to non-fossil fuels, all contribute to climate change and are important components on curbing carbon emissions. Nevertheless, constraints of carbon emission reductions are linked to economic growth and development (e.g., Schmalensee et al., 1998; Heil and Selden, 2001), therefore, emission (in)equalities are of key concern to, *inter alia*, climate negotiations, the design of mitigation policies, R&D investments and production reallocation decisions. This requires a deep understanding of the sources, driving forces and structure of emission inequality.

Over the last fifty years US cumulative CO_2 emissions exceeded the amount of a quarter billion tonnes, corresponding to an average share of approx. 23% of global emissions. Looking at the last decade, this figure dropped to 16% and 2017 was a 25 year low with slightly higher than 5 billion tonnes of CO_2 emitted in the atmosphere. However, despite the momentum in apparent technological advances and declines in renewable energy costs, 2018 CO_2 emissions from the US grew by 2.6% while the growth rate of total world was 2%. This does not compare favorably with the OECD (EU) countries CO_2 emissions which increased by 0.4% (decreased by 2%). Thus, the process of ensuring fair responsibility in global emissions abatement is agitated by the unequal distribution of emissions, as well as other disparities arising; such as the diverse speeds of economic development around the globe. Overall, the US have been the biggest CO_2 emitter since 1965, overtook only by China from 2004 onwards; even though 2004-2008 were all time zenith with US emissions exceeding 5.77 billion tonnes per annum.

As such, the US energy and environmental policies have become pressing matters in the global political arena, as emission reduction responsibility has both pivotal economic and social welfare implications. Further impediments came with the announcement in 2017 of the US withdrawal from the Paris agreement, in effect from late 2020 onwards. Therefore, the complexity involved in reducing US CO_2 requires the participation and coordination of all the states in the country. Even though in some cases individual states and/or broader regions themselves may intercede to provide incentives for energy efficiency and low carbon energies, it seems unlikely that systematic and persistent reductions will occur in the absence of a unified federal policy.

Regional heterogeneities can lead to diverse interests, agendas and/or incompatible perceptions about the fair distribution of the burden of emissions, all of which, may hinder different administrations to agree on commitments (Padilla and Duro, 2013). Understanding carbon inequalities across states is key for emission reduction policies as their patterns provide information about the underlying reasons contributing to the disparities and this might lead to differentiated state emission-reduction policies. Several factors have evolved differently across the US and the states do not follow a homogeneous structure as disparities in income, emissions, energy mix and intensity, production/consumption structure and energy efficiency, as well as conflicting political views with respect to environmental strategies, vary greatly. For instance, the bulk of domestic oil, gas and coal production originates from just a handful of states. As another example, consumption and production patterns present diversities linked not only to population size and structure (Zhu and Peng, 2012; Ramuzgo and Sarabia, 2015) but climate factors as well (York et al., 2003). Analysis of regional disparities in emissions per capita, and the factors that drive them, can provide valuable insights for establishing tailor-made mitigation policies which constitutes the understanding of the factors explaining the unequal distribution of emissions vital. The time-evolution of inequalities and the role of the driving factors, whether disparities concentrate between groups of states etc., deserve the attention of decision makers in order to achieve better policy directives. Inequalities in carbon emissions hinder that some states' emission-reduction potential is not fully exploited, e.g., through technology spillover effects or inter-state cooperation or federal coordination. High emission per capita disparity reveals the urgency of mitigating carbon emissions through knowledge and technology transfers from the more technology advanced and energy efficient states. For example, it would be more effective and fair to establish differentiated targets when the underlying pattern of inequalities is identified and their sources are known; e.g., these inequalities might be attributed to divergences in income per capita (e.g., Feng et al., 2018; 2021), the particular mix of energy sources, differing energy efficiency level, etc. Proper consideration of all these is necessary for the success of mitigation proposals.

The purpose of this paper is to assess the driving forces of US CO_2 emissions per capita, as well as their spatial and temporal attributes. To this end, we aim to provide two main contributions. First, we postulate one intuitively compelling approach to examine carbon emissions, based on the Kaya (1989) identity and Theil (1967) index as a reference.[1] Multiplicative factors can be used to analyze the distribution of per capita CO_2 emissions, explore carbon inequality and polarization, providing information about each factor contribution to inequality. Without doubt, systematic analysis of emissions inequality is useful to inform the decision-making process on mitigation proposals and activities by revealing the status of the relative responsibilities of different states/regions and the problems and causes associated with emission inequalities. Our second contribution is to clarify the nature of knowledge relative to US emissions, carrying out an empirical illustration over the period 1980–2017 for the 48 contiguous US states. In doing so, we use the concepts of inequality and polarization for different taxonomies of US states; based on the geographical location, geological structure, climate features and human development[2] indicators that might differentiate the attributes of a cluster of states over that of others. The empirical results present a comprehensive picture of US emission inequality and polarization to policymakers, and this way, we aim to advance knowledge regarding interrelationships among states broadly while also helping to inform regulators and decisions of environmental policies.

Findings reveal that inequality in CO_2 emissions increased between 1980 and 2017 while the peak figure is in 2011. Energy intensity, energy mix and labour productivity are the key inequality components while carbon intensity of energy fossil fuel use and employment are the least important. Therefore, policy measures focusing on either reducing the cost or increasing the efficiency of converting energy to GDP prove effective in controlling emissions, as convergence of energy intensity leads to a corresponding reduction in total CO_2 per capita inequality. Geographical grouping confirms that the set of measures to limit the concentration of CO_2 in the

[1] Note that our contribution is not the posit of an index decomposition methodology or the implementation of alternative index decomposition methods such us, e.g. logarithmic mean divisia index (LMDI). Theil decomposition with the help of Kaya and LMDI is just one method able to identify the driving factors that contribute to carbon emissions inequality. The importance of this method is highlighted by the fact that they offer simplicity in implementation and high decomposability (Ang, 2004). Previous research indicates that the particular methodology has become the norm in research similar to ours (e.g., Duro and Padilla, 2006; Padilla and Duro, 2013; Ramuzgo and Sarabia, 2015). However, recent studies (e.g., Roux et al., 2021; Roux and Plank, 2022) recommend alternative decomposition methods that claim to reflect more appropriately the common understanding of structure effects.

[2] Among others, Zaman et al. (2016) found a feedback relationship between CO2 and human development. It is worth noting that human development constitutes a key objective of economic growth. Energy utilization and intensity have an impact on human development as energy is an important resource for all economic activities. For example, there are several studies that highlight the link between energy consumption and human development indicators (e.g., Akizu-Gardoki et al., 2018). Moreover, generalizing the results of Greenstone and Hanna (2014), high human development may result in the capacity to fight pollution and increase effectiveness of environmental regulations. As such, it would be prudent to consider human development when drafting climate change mitigation policies.

atmosphere has amplified the inequality between PADD districts or Census divisions further than the within geographical location inequality. Using bipolar and multipolar groups, we report that between-group inequality contribution is rather high and within the range of 63-90%, on average across years; implying high heterogeneity between groups and high homogeneity within groups.

The structure of this paper is as follows. In Section 2, we briefly explain the background concepts of inequality and polarization. Section 3 discusses the econometric methodology, while section 4 describes the balanced panel of data employed and the construction of some key variables. Section 5 presents the empirical results and evaluates inequality across years, within and between groups of states. Finally, Section 6 concludes.

2. Background

A broad literature has emerged on the analysis of the distribution of income, its development over time, the identification of factors affecting inequality and their influence mechanisms, such as, to name a few, Gottschalk and Smeeding, (1997), Li et al. (1998), Forbes (2000), Acemoglu and Ventura (2002), Lin and Tomaskovic-Devey (2013), Madsen et al. (2018). Pronounced inequalities and their underlying causes are issues that need careful consideration when implementing mitigation initiatives, therefore, it is interesting to examine the distribution of CO_2 and the factors that explain it.

The notions of emissions inequality and polarization[3] are aimed to inform different emission abatement proposals on the contribution of each state to the climate change debate. Furthermore, economists have also noted the potential for income inequality to affect pollution indirectly through either the distribution of political power (Torras and Boyce, 1998) or changes in consumption (e.g., Ravallion et al., 2000), while the relationship between U.S. state-level CO2 emissions and income inequality was examined by Jorgenson et al. (2017).

An inequality index, which summarizes this distribution and used in this paper is the Theil (1967, 1972) index; the fact that it can be decomposed into different factors, makes it an appealing

[3] Polarization aims at examining the degree to which per capita CO_2 emissions are allocated across various states, but is different than inequality, as it is directly linked to groups of states presenting certain attributes and the degree to which antagonistic behaviors might form. Polarization is relevant with regard to reaching nation-level agreements, e.g., mitigation policies. For example, in the presence of strong internal group cohesion and certain group size, it is possible that some states will assume a dominant role and, therefore, undermine any bargaining process if their interests are not aligned with the nation’s behavior towards pollution.

candidate for this purpose. In the past decades, much research has analyzed international differences in CO2 emissions per capita via indicators of inequality, e.g., Theil, Atkinson and Gini (see Hedenus and Azar, 2005; Ramuzgo and Sarabia, 2015; Groot, 2010). For example, Heil and Wodon (1997, 2000) - of the earliest studies applied to energy and CO2 emissions inequality - used the Gini index to study the distribution of per-capita CO2 emissions and the impact of the Kyoto Protocol. Alcantara and Duro (2004) suggested the use of the Theil index. To the authors' knowledge, existing studies focus on international inequalities on a worldwide level, across OECD countries, or the EU (Padilla and Duro, 2013), therefore, this study constitutes the first one to exclusively analyze US states and broader regions.

In the literature, the determinants and sources of CO_2 inequality have been explored along various dimensions. For example, on the basis of the Kaya identity decomposition and CO_2 inequality determinants, Remuzgo and Sarabia (2015) find that international inequality was mainly caused by economic growth in terms of labor productivity while they argue that technology transfer, the type of transfer and its effectiveness allow a greater convergence in this factor and therefore in the CO_2 emissions. Similarly, Duro and Padilla (2006) also attribute inequality in per capita emissions to affluence while they also note that carbon intensity of energy and energy intensity contributions to inequality are not to be ignored.

On the other hand, assessments of the sources of CO_2 inequality have also emerged in different contexts. Duro and Padilla (2006) and Padilla and Duro (2013) examine the between- and within- country group inequalities in per capita CO2 emissions for global economies and EU countries, and report that the between-group component was the major contributor to the overall inequality. Moreover, Grunewald et al. (2014), find that 23% of the inequality in 2008 was sourced from natural gas use. The authors document reduction in the CO2 emissions global inequality and attribute this to the declining use of coal/peat and oil as well as the declining shares of emissions from the manufacturing and construction sectors.

Both the sources and determinants of inequality and both the spatial and temporal structure of inequality are relevant in understanding the distribution of emissions. CO_2 inequality and the extent to which it is influenced by several energy-related, macroeconomic, demographic or geology/climate related factors evokes the interest of this study, since persistently high and rising inequalities threaten socio-political stability and sustainability.

3. Methods

There are several standard measures of inequality in the literature. We focus on Theil (1967, 1972) index which fulfils the most widely accepted axioms including decomposability (see Cowell, 2000; 2011). The Theil measure belongs to the generalized entropy family of indices (Shorrocks, 1980, 1984). In particular, suppose N economies/regions under consideration and denote $y_i, \bar{y}$ the per capita CO_2 emissions of region *i* and global mean of y_i, respectively. If $p_i$the population share of the economy/region *i* in the total population., $\frac{y_i}{\bar{y}}$, then

$$T = \sum_{i=1}^{N} p_i \ln\left(\frac{\bar{y}}{y_i}\right) \quad (1)$$

As Bourguignon (1979) showed, Theil index has many desirable properties. It is additively decomposable in the sense formalized by Shorrocks (*1980, 1984*) and is transfer sensitive (Shorrocks and Foster, 1987), i.e., more sensitive to transfers at the bottom end of the distribution than at the top. This measure has been widely used to examine the evolution of income inequality (e.g., see Bourguignon and Morrisson, 2002). Theil entropy index treats equally differences in all parts of the distribution; alternative entropy indices such as the mean log deviation are more sensitive to changes at the bottom tail (Sarabia *et al.*, 2017). It has been very popular due to its relatively higher decomposability compared to others such as Gini coefficient or Atkinson index (Grunewald et al., 2014; Wang and Zhou, 2018). For example, Gini has been found to be over-sensitive to changes in the middle of the distribution and insensitive to changes at its top and bottom (Grunewald et al., 2014). In a nutshell, Theil index is the only population weighted inequality index that can be broken down into groups of observations, is differentiable, symmetric, invariant with scale and satisfies the Pigou-Dalton criterion. That is, *ceteris paribus*, any transfer of a region with, say, high level of emissions in our case, to another with a lower level should reduce or at least not increase the value of the index.

Further, to investigate the sources of inequality in CO_2 emissions per capita we use the Kaya (1989) identity as a reference. The latter, is a simplified yet successful model that evaluates emission drivers with economic, demographic and environmental factors (Tavakoli, 2018), i.e., the effect of most important-effective driving forces; selected on the basis of whether they can be reasonably thought to have an impact. The objective of this paper is to analyze CO_2 emissions per capita which summarizes important information about the asymmetries in emissions across the US. Kaya identity has found applications within several settings. To name a few, we note the use of Kaya by the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports; carbon emissions regional analyses of trends and their demographic, economic and technological driving forces (recent studies include Bianco et al., 2019, Wang et al., 2020, Patiño et al., 2021; De La Peña et al., 2022); carbon emissions sectoral analyses (e.g., Hammond and Norman, 2012 for an application to the UK's manufacturing sector or Eskander and Nitschke, 2021 for the UK universities progress in greening their energy sources); or even exploring thermal energy use in buildings (Ürge-Vorsatz et al., 2015 or Mavromatidis et al., 2016). Although we acknowledge that Kaya identity cannot accommodate the full complexity of the cause-effect relationships among the drivers, the identity postulates a comprehensive framework to understand these drivers (e.g., Bianco et al., 2019).

This way, Theil index can be decomposed to inequality attributable to different factors. CO_2 emissions per capita are decomposed into the product of five factors: (i) carbon intensity of fossil fuel use, the ratio of CO_2 emissions to fossil fuel consumption (FC); (ii) energy mix, the ratio of FC to total primary energy consumption (EC); (iii) energy intensity, energy used as a percentage of real gross domestic product (GDP); (iv) labour productivity, GDP per worker; (v) employment rate; working over total population. Mathematically, this can be expressed as

$$y_i = \frac{CO_{2,i}}{TP_i} = \frac{CO_{2,i}}{FC_i} \cdot \frac{FC_i}{EC_i} \cdot \frac{EC_i}{GDP_i} \cdot \frac{GDP_i}{EP_i} \cdot \frac{EP_i}{TP_i} = a_i \cdot b_i \cdot c_i \cdot d_i \cdot e_i \tag{2}$$

where, in our study, *i* represents a region, i.e., state of the US. Next, define four hypothetical CO_2 emissions per capita vectors for each state *i* by permitting in each vector only the value of one factor f divergence from the global mean

$$y_i^f = \frac{f_i}{\bar{f}} \left(\bar{a} \cdot \overline{b} \cdot \bar{c} \cdot \bar{d} \cdot \bar{e}\right); for\ f = a, b, c, d, e \tag{3}$$

Let $\bar{y}^f = \sum_{i=1}^{N} p_i y_i^f$. To measure each factor's contribution to the global inequality index, the Theil index is decomposed as

$$I^f = \sum_{i=1}^{N} p_i \ln\left(\frac{\bar{y}^f}{y_i^f}\right) ; f = a, b, c, d, e \tag{4}$$

That is, each index measures the partial contribution of each factor to global inequality. The corresponding Theil indices for these factors using the global average of CO_2 per capita as a reference

$$T^f = I^f + \sum_{i=1}^{N} \ln\left(\frac{\bar{y}}{\bar{y}^f}\right) = \sum_{i=1}^{N} p_i \ln\left(\frac{\bar{y}}{y_i^f}\right); f = a, b, c, d \tag{5}$$

$$T = T^a + T^b + T^c + T^d + T^e = \sum_{i=1}^{5} I^f + \sum_{i=1}^{4} \ln\left(\frac{\bar{y}}{y_i^f}\right) \tag{6}$$

The second term of the Eq. (6) measures the interaction between the different factors considered; that is, they are the factorial correlations between: (i) carbon intensity of fossil fuel use (α) and fossil fuel consumption per capita ($bcde$), i.e., $\ln\left(\frac{\bar{y}}{y_i^a}\right) = \ln\left(1 + \frac{\sigma_{\alpha,bcde}}{\bar{y}^a}\right)$; (ii) energy mix ($b$) and energy consumption per capita (cde), $\ln\left(\frac{\bar{y}}{y_i^b}\right) = \ln\left(1 + \frac{\alpha\sigma_{b,cde}}{\bar{y}^b}\right)$, (iii) energy intensity (c) and GDP per capita (de), $\ln\left(\frac{\bar{y}}{y_i^c}\right) = \ln\left(1 + \frac{\alpha b\sigma_{c,de}}{\bar{y}^c}\right)$ and (iv) labour productivity (d) and employment rate (e), $\ln\left(\frac{\bar{y}}{y_i^d}\right) = \ln\left(1 + \frac{\alpha bc\sigma_{d,e}}{\bar{y}^d}\right)$; where $\sigma_{\alpha,bcde} = \sum_{i=1}^{N} p_i(\, a_i - \bar{a})\left(b_i \cdot c_i \cdot d_i \cdot e_i - \overline{bcde}\right)$ is the population weighted covariance between variables a_i and $b_i \cdot c_i \cdot d_i \cdot e_i$. Therefore, apart from inequality in emissions per capita attributable to the five factors, we also account for inequality attributable to the interaction terms (see, e.g., Duro and Padilla, 2006).

Suppose now that the different regions may be grouped consistently with some vector of attributes, such that member-states demonstrate similarities, yet different clusters of states present unrelated attributes. For example, Wolfson (1994) constitutes one of the earliest developments of polarization measures, designed to capture the disappearance of the middle class and analyzing the bipolar case. The seminal work of Esteban and Ray (1994) on multiple-pole cases formulated polarization indices based on a behavioural model and the identification-alienation nexus (see also Duclos et al., 2004). In this regard, polarization increases when there is strong cohesion within-group and long distance between the groups, i.e., within-group identity vs. between-group alienation can promote conflict. Zhang and Kanbur (2001) suggested a polarization measure, based on the inequality decomposition by groups (Shorrocks 1980, 1984). In these lines, let e_g the share of the region g in the US population, T_g the inequality in the region g, $\bar{y}^g$ the average CO_2

emissions per capita in the region g and, finally, G the number of regions. The decomposition of the total inequality in the between- and within-group components can be expressed as

$$T = \sum_{g=1}^{G} e_g T_g + \sum_{g=1}^{G} e_g \ln\left(\frac{\bar{y}}{\bar{y}g}\right) = T_w + T_B \quad (7)$$

where T_w is the within-group inequality component representing the spread of the distributions in the individual subgroups, T_B is the between-group inequality component representing the spread between the group means.

Therefore, the ratio T_w/T_B can be regarded as a scalar polarization index, henceforth ZK (Zhang and Kanbur, 2001). This captures the average distance between the groups in relation to the per capita CO_2 differences within groups. More homogenous groups would imply higher polarization, as internal cohesion will tend to magnify the differences across groups, i.e., T_W and T_B can be perceived as measures of group identification and alienation. Finally, note that with this approach, T_w and T_B expressions can be further explored by carrying out a factor decomposition analysis (e.g., see Duro, 2010).

4. Data and preliminary analysis

This study is based on US total CO_2 emissions in million metric tonnes calculated from fossil fuel combustion in the residential, commercial, industrial, transportation and electric power sectors; obtained from the Energy Information Administration (EIA) (https://www.eia.gov). For the Kaya identity factors (Section 3.1 and Eq. 2), fossil fuel consumption and total primary energy consumption are also obtained from EIA while state-level GDP, employment and population data are obtained from the US Bureau of Labor Statistics. To convert nominal GDPs to real, in the absence of state-level data, we use the regional Consumer Price Indices from the Bureau of Labor Statistics (West, South, Northeast and Midwest CPI). Table 1 (Panel A) lists all the relevant variables. The list of data used and their sources is provided in Table 1. The dataset covers the 48 contiguous US states with annual observations comprising a balanced panel of 48 regions with time span from 1980 to 2017.

[INSERT TABLE 1]

Figure 1 shows the geographical structure of the US in terms CO_2 per capita and Kaya factor intensities. Each row of the plot reflects the deviations of either CO_2/P, CO_2/FC, FC/EC, EC/GDP, GDP/P or EP/P from the corresponding national level in 1980 (left subplots) and 2017 (right subplots); 1 implies that state-variable is equal to the US population weighted average. Taxonomy is based on pre-defined thresholds from the distribution of the 1980 deviations. In particular, white (black) indicate the states that have recorded relatively low (high) values CO_2/P and Kaya factors; these correspond to the lower and upper decile, respectively. Grey is used for the remaining states with the light (dark) tone indicating the lower (higher) quartile. Note that for 2017 we also use the 1980 thresholds in order to identify changes throughout 1980-2017. For example, the top-row colorbar is {< 0.6, 0.7, 1.2, > 1.7}, i.e., the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles of the 1980 distribution of (State-CO_2/P)/(US-CO_2/P) where US-CO_2/P is proxied by the population weighted average of CO_2/P across the US. E.g., Mississippi (dark grey, 2017) recorded 20%-70% higher CO_2/P, compared to the US average; emitting more on a per capita basis relative to 1980 which was within the bounds -30% to +20% of the average.

[INSERT FIGURE 1]

The high emitters in 1980, Louisiana, Wyoming, West Virginia and North Dakota, maintained their status in 2017; more than 70% per capita emissions compared to the average-state. It is worth noting that these states exceed US-CO_2/P by more than 2.65 (4.35) times in 1980 (2017). Texas reduced CO_2/P, from 85% in 1980 to 58% in 2017 above the US average. In 2017 the list is augmented by Indiana and Montana. Similarly, low per capita emitters in 1980, New York, Oregon, Rhode Island and Vermont, still emit less than 60% of the amount an average US citizen does. An exception is Idaho (56% in 1980 vs. 68% in 2017), replaced though by California, Connecticut, Massachusetts and Maryland; the latter being the most notable change, from 76% in 1980 to 54% in 2017. In 2017, the highest (lowest) emitting state, on a per capita basis was Wyoming (New York) with (State-CO_2/P)/(US-CO_2/P) more than 6.5 (close to 0.5). This implies that the average person in Wyoming emits 13 times more compared to the typical New Yorker! However, in absolute CO_2 terms, NY in 2017 emitted 2.5 times more than Wyoming.

We next turn to Kaya factors. First, the carbonization index (CO_2/FC) was relatively high in 1980 for states such as Wyoming, West Virginia and North Dakota, i.e., CO_2/FC more than 15% compared to the national 1980 average. This also holds for Nebraska and North Carolina, but for them, in 2017, CO_2/FC shifted close to the average US figure. However, Kentucky, Montana and Missouri are added to the high CO_2/FC (>1.15) states. Moreover, persistently low CO_2/FC states are Louisiana and Texas (close to 86%-90%). Yet, Arkansas, California and Oklahoma no longer belong to the CO_2/FC<90% of the US 2017 average group and Connecticut, Delaware New Jersey, Rhode Island and NY take their place.

Second, high fossil fuel portion in the energy consumption mix (>1.1 of the national average) describes best Delaware, Louisiana, New Mexico and Texas. It seems that in 2017, Kansas diversified its mix, reducing this ratio during 1980-2017, from 11% higher to 11% lower than the national average. Still, quite a few more states (ten) have now FC/EC ratio more than 10% when compared to the US average. In 2017, of the more diversified states are Maine, Oregon, Washington, New Hampshire and South Carolina; while the former three were also part of the 1980 list of low FC/EC (together with Vermont and Idaho).

Third, high energy intensity states in 1980 are Alabama, Louisiana, Montana, West Virginia and Wyoming, with EC/GDP more than 90% above total US intensity, on average. In 2017, these states still face high cost of converting energy to GDP in 2017; EC/GDP higher than 60% above the US average together with Arkansas, Iowa, Mississippi, New Hampshire, North Dakota and Oklahoma. 2017 EC/GDP of these high intensity states is approx. 2.5 times higher than that of the national level. Then again, for California, Connecticut, Massachusetts, NY and Rhode Island the implied cost of converting energy to GDP is lowest; 30% lower than the national average. In 2017, we add Delaware, Maryland and New Jersey with energy intensities lower than 70% of the total US EC/GDP. It is worth noting that as of 2017, NY and Massachusetts constitute the most efficient states in terms of converting energy to GDP, with figures close to 43% of US EC/GDP followed by California and Maryland (close to 50%).

Fourth, of the high affluence states, it is only NY which remains at the top throughout 1980-2017, improving the index of GDP/EP from 14% above the US average to almost 40%. Illinois, Louisiana, Texas and Wyoming, although highly affluent in 1980, they now report a ratio of (State- GDP/EP)/(US- GDP/EP) less than 1.1. The most pronounced difference is Louisiana whereby from GDP/EP above 40% of the average US 1980 figure, in 2017 it was 5% below. On the other hand, in 2017, California, Connecticut, Delaware, New Jersey, Washington, and Massachusetts all exhibit GDP/EP at least 10% above the US average affluence. Notable example is Massachusetts; value of approx. 10% below (20% above) average US GDP/EP in 1980 (2017).

Finally, with regards to employment as share of state population, top are Colorado, Minnesota, New Hampshire, North Dakota and Connecticut in 1980 with EP/P at least 10% above the US average employment. Of these states, the former three still appear in the 2017 list together with Iowa, Massachusetts, North Dakota, Vermont and Wisconsin. In 1980, states that faced employment figures lower than 90% of the US average are Alabama, Louisiana, Mississippi, New Mexico and West Virginia. The overall picture remains the same in 2017 with the addition of Arkansas; though the latter is marginally close to 90% in both years.

5. Empirical Results

Table 2 shows the factorial decomposition of US states' inequality in CO_2 emissions per capita using the Theil index. We report all years from 1980 through 2017. Findings reveal that inequality in CO_2 emissions increased by 38% between 1980 and 2017 while the peak figure is in 2011 with a Theil index of 0.1112. After 2011, CO_2 per capita inequalities show negative y-o-y changes apart from 2013 and 2017 where annual growth rates stood at 0.26% and 3.87%, respectively.[4] Post-2005 period Theil indices are above the 0.0935 average across years, implying that responsibilities for CO_2 have not diffused in the last decade. The most pronounced increase (decrease) occurred between 1989-1990 (2014-2015), i.e., an upsurge (down surge) in excess of 9% (-6.9%). Still, in the latter case the national level of inequality was above average.

[INSERT TABLE 2]

Energy intensity T^c has an average contribution close to 99%; 90% confidence interval (CI) of 76.9-114% across years. These results contradict Duro and Padilla (2006) 3-factor and

[4] Between 2011 and 2017, coal consumption and production in the US dropped by more than 29%. This continuous decline coincides with the shale gas boom after 2008 which made natural gas a baseload fuel and pushing coal lower in the electricity generation mix. Note also that in 2013 (2017) coal consumption (production) increased by 3.8% (6.6%) relative to the previous year. Figures are obtained from Energy Information Administration (EIA).

Remuzgo and Sarabia (2005) 4-factor decomposition (at a global level though), who find that affluence (GDP per capita or by active population) is the major factor in explaining international CO_2 inequalities during 1971-1999 and 1990-2010, respectively. Still, the next key inequality components are energy mix T^b and labour productivity T^d with partial contributions above 8.5%. Carbon intensity of energy fossil fuel use T^a and employment rates T^e are the least important, in line with Remuzgo and Sarabia (2005).

Therefore, policy measures focusing on either reducing the cost or increasing the efficiency of converting energy to GDP prove effective in controlling emissions, as convergence of energy intensity leads to a corresponding reduction in total CO_2 per capita inequality. Indicative strategies might include incentives to high intensity states for the development/use of advanced technologies in energy conversion, technology transfers, allocating the production of certain manufactured products to low energy intensity regions/states (rather than producing them in-state) or even infrastructure investments to facilitate use of fuel efficient vehicles, mass transportation and carpools. The contribution of *EC/GDP* strengthened by 40 (18) percentage points relative to its level in 1980 (2010) while inequality in the energy intensity increased by about 113%, from 0.0557 to 0.1188 (T^c). A possible reason might be, *inter alia*, the prevalence of more pronounced asymmetries in energy-related technological developments across states.

Affluence, after 1993 also exhibited an upward path contributing to the increasing Theil indices. In 2017, *GDP/EP* inequality is up by more than 69% compared to the average and its contribution is maximized at 13.4% in 2017. High affluence is translated to more disposable income to spend on, among others, (i) electricity - not just for essential lighting, heating and refrigeration but also for air conditioning and gadgets, and (ii) gasoline for larger cars - especially SUVs. On the industrial side, high affluence is associated with increased demand for petroleum products which also translates into bigger throughputs for refineries, which are one of the main polluters. Thus, convergence in affluence/productivities of the employees is necessary to limit CO_2 inequalities across states which might apply to policies targeting at diffusing new technologies, equipment and promoting technology transfers across states from high to low productivity states, and changes in the work models (adopting technologies) that lead to more efficient human work and energy use.

On the contrary, energy mix contribution has noted an overall decay of its relative importance mainly after 1983 while T^b is down by 13% (19%) since 1980 (1990). This is indicative of the progress that has been made after the 1970s oil price crises where most regions now use a relatively more diversified energy matrix, including renewables and natural gas; condensing the weight of this factor throughout the years. However, energy mix still it holds its relative importance above 6.5% and strategies that promote diversified energy portfolios across states are still expected to lower CO_2 inequalities within the US. Furthermore, CO_2/FC partial contribution is at a maximum in 1999 while, on average, it has caused more than 3.6% of total inequalities. Since 1999 its weight has decreased to less than 2.5% reflecting the gradual adoption of cleaner energy or cleaner technologies across the US. Finally, *EP/P* has average contribution of 1.6% and is consistently above average only during the 1980s. High employment is typically associated with economic growth where energy use is intensified.

Table 2 presents also the contributions of the interaction terms. The partial contribution to national CO_2 inequality of the interaction between energy mix and energy intensity, i.e., *inter*(*FC/EC, EC/GDP*) is 13.1% on average with a consistently positive contribution. Heavy energy consuming regions tend to include higher portion of fossil fuels in their consumption mix, amplifying cross-state inequalities.

On the other hand, *inter*(*EC/GDP*, *GDP/P*) has contributed negatively with an average value of 36.6% in absolute terms. This negative correlation means that, *ceteris paribus*, less (more) affluent regions face a relatively higher (lower) cost of converting GDP to energy; due to less available funds for R&D, infrastructure projects' (related to efficient energy use) delays and/or slower rates of adoption of new energy smart technologies, among others. That is, high emitting rich states exhibit lower energy intensities, which reduces in turn CO_2 inequalities (see also Duro and Padilla, 2006).

In addition, the contribution of *inter*(CO_2/FC, *FC/P*), has steadily increased (90% CI of approx. -/+ 9%).[5] Pre-1995, states with greater *FC/P* emitted less CO_2 per unit of fossil fuel consumed. This could be attributed to relatively higher economic growth pre-1995, a diverse degree of industrialization across states and consumer behaviour (heavy fossil fuels consumers are likely to adopt more progressive technologies) which might have had a balancing effect on inequalities. However, post-1995 this dynamic interaction shifted with high *FC/P* states emitting

[5] Since 1980 the US fossil fuel consumption has increased by approx. 13% due to economic growth and increased demand for energy, as well as demographics; CO_2 have also increased by 7.3% relative to their 1980 level. However, with the gradual adoption of new cleaner and more efficient technologies that are able to reduce carbonization indices, nation-level CO_2/FC has experienced a decrease of more than 5% whereas *FC/P* a decrease of more than 20%.

more per unit of fossil fuel, thus, magnifying total CO_2 inequalities. Finally, *inter*(*GDP/EP*, *EP/P*) has a limited contribution within a 90% CI of -/+ 1.5%. Pre-2008, high employment was associated with lower labour productivity; partially attributed to diminishing marginal returns due to the nonlinear relationship between workforce and productivity. However, in the aftermath of the 2008-2009 financial crisis this contribution becomes positive highlighting the impacts on employment[6] and economic growth.

5.1 Inequalities within and between groups

Tables 3-6 summarize the findings obtained from the decomposition of US CO_2 emissions per capita within and between-group inequality components using the Theil index. For the classification of states to groups, we use both geographical data (Table 3, Panel A) and key coding variables (Table 3, Panel B and Tables 4-6).

Geographical partitioning follows the Census Bureau categorization to four regions or nine divisions, i.e., New England and Mid-Atlantic (Northeast); East and West North Central (Midwest); South Atlantic, East and West South Central (South); Mountain and Pacific (West).[7] For robustness we also consider the five PADD districts (EIA): PADD1 (New England, Mid & South Atlantic), PADD2 (East & West North Central plus Kentucky, Oklahoma, Tennessee), PADD3 is the Gulf Coast (Alabama, Arkansas, Louisiana, Mississippi, New Mexico, Texas), PADD4 the Rocky Mountain (Colorado, Idaho, Montana, Utah, Wyoming) and PADD5 the West Coast (Pacific plus Arizona, Nevada).

Groups are also formed on the basis of key coding variables including state-level CO_2 per capita, energy production vs. consumption, geology of states (e.g., coastline length), climate (e.g., Palmer's drought severity index), human development indicators (e.g., education attainment). For this task, we set an ad hoc minimum group size to ten states while for simplicity, each group

[6]The number of people employed people in the US was gradually growing during 1980-2008 with average annual growth rate of approx. 1.4% (this rate is faster than population increase in the US of approx. 1% over the same period or even the increase in the working population which stands higher at 1.3%) while the year 2009 there was a more than 3.5% decrease in the figure, amplified by a further drop of approx. 0.13% in 2010.

[7] The state allocation to divisions is: New England (Connecticut, Maine, Massachusetts, N. Hampshire, Rhode Island, Vermont), Mid Atlantic (N. Jersey, N. York, Penns.), East North Central (Illinois, Indiana, Michigan, Ohio, Wisconsin), West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, N. Dakota, S. Dakota), South Atlantic (Delaware, Florida, Georgia, Maryland, N. Carolina, S. Carolina, Virginia, DC, W. Virginia), East South Central (Alabama, Kentucky, Mississippi, Tennessee), West South Central (Arkansas, Louisiana, Oklahoma, Texas), Mountain (Arizona, Colorado, Idaho, Montana, Nevada, N. Mexico, Utah, Wyoming) and Pacific (Alaska, California, Hawaii, Oregon, Washington). Note though that Alaska, Hawaii and DC are excluded from our analysis.

includes an even number of states, i.e., 10, 12, 14, …., S states. It follows that, the maximum group size for bipolar groups, tripolar and quadpolar groups is S = 38, 28, 18; leading to 15, 55, and 35 groups, respectively.[8] Consider CO_2/P as the variable of interest; thus the opportunity set of groups sizes is $\{(10,38)_1, (12,36)_2, \dots(24,24)_8, \dots, (36,12)_{14}, (38,10)_{15}\}$. For example, $(12,36)_2$ splits states to: emitters that belong to the top quartile of the distribution (12 states) vs. others (36). Likewise, a group size for $(36,12)_{14}$ defines emitters that belong to the low quartile of the CO_2/P distribution vs. others. $(24,24)_8$ splits the distribution of CO_2/P in half (median, i.e., 24 high vs 24 low emitters), and so on.

[INSERT TABLE 3]

[INSERT TABLE 4]

For brevity, Tables 3-4 report the average % contributions to total inequality across 1980-2017. Theil indices in absolute terms are not reported and are available from the authors upon request. The estimation of the within/between group inequalities renders a wealth of results, from which we choose those that are most pertinent to the main objective of the paper.

5.1.1 How much inequality can group dynamics explain?

Results when groups are based on geographical partitioning are summarized in Table 3, Panel A. First, between and within components contributed to the overall inequality during 1980-2017, 28-72% (Census regions), 60-40% (Census divisions) and 55-45% (PADD districts), respectively, on average. Between-group figure was lower than within-group inequality contribution in the four-regions, confirming the importance of appropriately reflecting state-level diversities beyond Northeast, Midwest, South and West regions. Second, the between-group contribution to total inequality has overall increased; especially for the PADD district grouping (32 bps per year); less so for the group dynamics in the nine divisions (13) and marginally zero for the broad regions (-1). This result, coupled with the general increase in Theil statistic (1980-2017, Table 2), provides evidence that, the set of measures to limit the concentration of CO_2 in the

[8] For instance take the quadpolar groups. Group sizes considered encompass 10, 12, 14, 16 and 18 states (5 possibilities). The number of permutations with repetition allowed and order important are $n^r = 5^4 = 625$ (choose 4 out of 5). However, conditional on the restriction that the sum of states should be 48, this results in 35 possibilities.

atmosphere has amplified the inequality between PADD districts or Census divisions further than the within geographical location inequality; this seems masked in the limited four-region case. Third, the growing importance of the between-group inequality implies that CO_2 emission management requires administrations and policy makers to consider group dynamics as well.

Moreover, the ZK (T_B/T_W) measure (Zhang and Kanbur, 2001), stands at 0.39 with a 90% CI of 0.34 to 0.45 for the Census regions while polarization seems to have fallen throughout the years at a rate of 0.06% pa. For the Census divisions and PADD districts, both figures stand above 1, i.e., 1.51 (90% CI of 1.25-1.68) and 1.22 (90% CI of 0.91-1.53), respectively, while polarization has experienced an increase of more than 0.5% pa. Overall, the latter two geographical consortia involve more homogenous groups than Census regions; internal cohesion is more pronounced for the Census division, while for PADD districts the contrast between homogeneity and heterogeneity is relatively balanced.

Next, the results for within and between group contributions to CO_2 inequality are presented in Table 3, Panel B. The reported figures are the groups (within each set of bipolar and multipolar groups) for which the between-group inequality component's explanatory capacity of total inequality is maximized, i.e., $\overline{ZK}$ is max. We also report equally sized groups for completeness, based on median, tertiles and quantiles of the CO_2/P distribution, i.e., 24, 16 and 12 states in each group for bipolar, tripolar and quadpolar, respectively. At the bottom of the table the results for a quintipolar group is also noted (minimum group size constraint is relaxed to 8 states).

Overall, we can see that between-group inequality contribution is rather high and within the range of 63-90%, on average across years; implying high heterogeneity between groups and high homogeneity within groups. Even for bipolar groups, ZK measure is higher than 1.7 in all cases with all 90% CIs no less than 1.45. For multipolar groups T_B/T is no less 80% with ZK higher than 4.2. In most cases we observe a slight annual decrease in the capacity of between-group inequality to explain total inequality (-3 to -10 bps pa on average); exception is the tripolar (based on tertiles) and bipolar (26 high vs 22 low emitters) grouping where an anemic growth factor of 3-5 bps pa is noted; reflected also in the annual growth factor of ZK (0.16-0.23% pa).

[INSERT TABLE 5]

Furthermore, we allocate states into bipolar (Table 4) and tripolar groups (Table 5) on the basis of Kaya identity factors' distribution. Energy intensity grouping (36 highest vs 12 lowest energy intensity states) explains more than 57% of per capita CO_2 inequalities. Similarly, splitting the 48 US states based on their energy mix, the explanatory capacity is rather lower, but still more than 50% (34 highest vs 14 lowest fossil fuel users; % of total state-level energy consumption).

As $T_B/T > 50\%$ for both energy intensity and mix groupings, $\overline{ZK}$ across 1980-2017 exceeds 1. Nevertheless, the negative annual growth rates of ZK signal a decrease in explanatory powers of the between-group components. If we introduce a third polar then the amount of inequality explained rises but not dramatically: a marginal improvement of more than 1.5% and 20%. Note that for quadpolar groups the improvement over tripolar is limited and less than 50 bps in both cases; results not presented here and are available from the authors upon request. The explanatory power of groups of states, formed on the basis of energy mix, still has negative trend, however for energy intensity groups the trend of the importance of tripolar and quadpolar groups increases by more than 8 bps per year; equivalent to a positive growth rate for the yearly ZK of more than 0.4%. Finally, carbon intensity, labor productivity and employment status groups explain inequalities by less than 20% while the ZK, within the range 0.09-0.23 is maximized for carbon intensity; explanatory power increases as reflected in the change (growth) of T_B (ZK).

Looking next at the remaining factors, the following observations are in order. First, primary (fossil fuel) energy balance, ie., production vs. consumption explanatory power stands at 48-60% (46-50%) albeit the importance of such factor(s) has(ve) decreased overall as reflected in growth rates of T_B and ZK. Average ZK ranges from 0.87-1.51 depending on the type of energy balance (primary vs fossil fuel) and size of groups. The figure for GDP from the manufacturing sector as a share of total GDP (population density) is 17-18% (36-42%) with ZK of roughly 0.23 (0.58-0.73).

Second, geology-based grouping, based on either mean elevation, forest cover or coastline/shoreline mileage explanatory power, stand between 5.5-more than 33% with most (least) important factor the forest cover (coastline/shoreline); $\overline{ZK}$ of 0.32-0.50 (<0.1). Third, climate-based allocation to groups, based on either average temperature, precipitation, drought, HDD and CDD explanatory power range within 3.9-29.3% with most important factor the drought severity index which $\overline{ZK}$ is 0.19-0.42; and least relevant factors both precipitation and HDD ($\overline{ZK} < 0.15$). Forth, as for the development-status groups (education attainment, life expectancy, real disposable income and human development) all support relatively high explanatory capacities (more than 37%), with education being the most prominent; $\overline{ZK}$ of 0.80-1.16.

Finally, forest cover, drought severity and education attainment constitute the most important factors in each block, i.e. geology, climate, development, respectively. It is only though the drought severity which shows an positive trend within 1980-2017; T_B (ZK) average increase of 15-34 bps (0.8-2.5%) pa.

5.1.2 Robustness checks and partial contributions

To obtain a benchmark on the potential of explaining inequalities approximate the distribution of T_B/T (or $\overline{ZK}$) by implementing permutations of the 48 states without repetition and randomly assigning states into groups. We use 3,000 random permutations of the 48 contiguous states. These are further then allocated based on 15 partitions of different group sizes, where minimum group size is no less than eight states i.e. $\{(10,38)_1, (12,36)_2, \ldots (36,12)_{14}, (38,10)_{15}\}$, resulting in 45,000 (=3000x15) random bipolar groups. In a similar way, this leads to 165,000 (=3000x55) and 105,000 (=3000x35) random tripolar and quadpolar groups, respectively; i.e. a total of 315,000 cases. Then, we calculate the confidence interval of T_B/T (or $\overline{ZK}$) seperately for bipolar, tripolar and quadpolar groups and compare this to the observed statistic of the exogenously set groups. Asterisks ***, **, * in Tables 4 and 5 indicate that the corresponding statistic (T_B/T or $\overline{ZK}$) of the defined grouping exceeds the right-side of the 99, 95, 90% confidence interval of T_B/T or $\overline{ZK}$ whereas numbers in bold compare this with the upper quartile of the simulated value.

Results in Tables 4 and 5 are consistent across T_B/T and $\overline{ZK}$. Collectively, we can see that it is only for the coastline/shoreline, precipitation index and HDD values that the observed T_B/T or $\overline{ZK}$ is less than the upper quantile of the simulated values. When considering multipolar groups we add to the list labor productivity and employment. Finally, from the segregation of states into groups we observe that, at least for the most important partitions - i.e., energy mix and intensity (Kaya), energy balance (other), forest cover (geology), drought severity (climate), education attainment (development indicators) – the reported T_B/T or $\overline{ZK}$ always exceeds the 95% right tail of the simulated ones.

[INSERT FIGURE 2]

Figure 2 shows additional information with respect to the contribution of each Kaya factor on T_B (left panel of the figure) and T_W (right panel). Results are overall consistent to our previous analysis in Table 2. More detailed information is also presented in Table 3 for location-based and state-level CO_2 based groups as well as Table 6 for selected cases; the groups that maximize ZK for each partition. In Table 3 it can be seen that, the most influential contribution to T_B and T_W is still energy intensity (as in Table 2) with an average individual contribution higher than 79% and 127% for T_B and T_W, respectively. This factor has increased its contribution throughout the years experiencing a steady momentum of more than 0.78% pa.

[INSERT TABLE 6]

For between-group inequalities the partial contribution to national CO_2 inequality of *inter*(*FC/EC, EC/GDP*) has been consistently the second most important factor with values of higher than 26.5%, yet decreasing throughout the years by 10-21bps pa. Overall, we can see the dominance of energy intensity in explaining inequalities for both between and within groups; although the effect is more pronounced for the former.

6. Discussion and Conclusions

The last fifty years have witnessed rapid economic development. Inevitably, with the increased energy demand that fuelled that growth, energy consumption and CO_2 emissions have both experienced a more than twofold increase worldwide. Increase in energy consumption in the US (OECD) stands close to 50% (75%). CO_2 emissions have increased in the US (OECD) by 23% (33%), while the US has been responsible for a share of 40-45% of OECD countries emissions throughout the years. Since carbon emission reductions are linked to economic growth and development (e.g., Heil and Selden, 2001), emission inequalities are of central interest to the design of mitigation policies.

This paper has analyzed the polarization of per capita CO_2 emissions in the US through the use of Z-K index (Zhang and Kanbur, 2001), whose main added value is its ability to be decomposed by factors. In this case, we have also proposed a multiplicative decomposition of this index by using the factors of Kaya (1989) to explore carbon inequality and polarization. The distribution of per capita CO_2 emissions uncovers the contribution of each factor to inequality and advances knowledge regarding interrelationships for different taxonomies of US states. The results of the factorial decomposition show that US inequality in CO_2 emissions per capita has increased by 38% between 1980 and 2017 while post-2011 inequality level shows negative trend, with main exception the year 2017; when annual coal production increased by more than 6.5%. Overall, post-2005 period Theil indices indicate that responsibilities for CO_2 have not diffused in the last decade.

When considering mitigation efforts, identification and assessment of the most influencing factors which control emissions can help devise effective policies to reduce emissions. It is observed that the bulk of inequality was caused by energy intensity while energy mix and labour productivity are also important. Thus, increasing the efficiency of converting energy to GDP allows a greater convergence in this factor and therefore in emissions. This is important for decision-makers when considering mitigation proposals since promoting technologies that reduce energy intensity can prove useful in controlling emissions. Such strategy requires the continued development of either new and enhanced technologies, improved industry access to technologies, technology transfers from one state to another and government provisions and incentives together with coordination among states with a view to reduce existing asymmetries and rationally allocate and utilize resources across states. For example, in 2017 high energy intensity states such as Louisiana and Wyoming have ratios of energy intensity above 2.5 times the US average, while New York and Massachusetts score at less than half the average. At the same time, it is crucial for the rational allocation and utilization of resources across states. Even as renewable technologies become a viable part of our energy future, the US carbon budget should be justified from climate policies that ensure its most equitable use.

That said, the discussion to achieve reduction of carbon emissions and the distribution of mitigation efforts among states is a current issue. It requires knowledge of the factors that determine the differences in emissions between states. We document a growing importance of the between-group inequality which implies that CO_2 emission management requires administrations and policy makers to consider group dynamics as well. The structure of different US states is not homogeneous; disparities in income, emissions, energy mix and intensity, production/consumption structure and energy efficiency, or even conflicting political views with respect to environmental strategies, vary greatly among states. However, for carbon taxes and cap and trade systems it is important to know who causes emissions and why. For example, it has been argued that climate mitigation policies need to take into account distributional effects on different income groups (Feng et al., 2018; 2021). Based on the geographical location, geological structure, climate features and human development indicators we find that the attributes of a cluster of states over that of others might differentiate. These differences and their driving forces have implications for the willingness to share the burden of emission mitigation within the US. For example, as most of domestic oil, gas and coal production originates from just a few states, regional heterogeneities can lead to different perceptions about the fair distribution of the burden of emissions and different agendas which can act as an obstacle to share objectives about targets and/or agreements.

Of course, this research is one of the first steps and although Kaya identity is a comprehensive tool that can be used for this evaluation, more complicated models can be employed for better understanding of states in the subject of carbon management. As this is the first study to comprehensively assess the per capita emissions across the US states as well as for groups of states (e.g., within and between groups, using geographical, geological, climatic and human development partitions of states), there is scope to potentially extend our analysis. For example, a separate but related research issue could be to focus on decomposing carbon emissions per unit of GDP (carbon intensity, e.g., Duro et al., 2016; Tian et al., 2021) to quantify the importance of the carbon driving forces rather than decomposing emissions per capita. In addition, although our decomposition follows an extended Kaya identity, there are other formulations conceivable perhaps equivalently able to offer further insights. For example, factoring out the employment rate factor and substituting it with labor force participation (working over adult population) and population structure (adult population/total population) might result in different interpretations. However, the postulated method can easily be generalized to accommodate the latter extension; it would lead to a more sophisticated Kaya equation but less parsimonious. Overall, several factors including, *inter alia*, fossil fuel intensity or per capita fossil fuel consumption might have confounding effects, which might be an interesting topic itself. In a nutshell, given the increasing emphasis on emission levels, asymmetries and environmental concerns, there is a proliferation of measures capturing different aspects. Creating analyses using alternative formulations in the modelling procedure, albeit an important research question, is left for future research.

References

Acemoglu, D., J. Ventura, J., 2002. The world income distribution. Quarterly Journal of Economics 117, 659-694.

Akizu-Gardoki, O., Bueno, G., Wiedmann, T., Lopez-Guede, J.M., Arto, I., Hernandez, P., Moran, D., 2018. Decoupling between human development and energy consumption within footprint accounts. Journal of Cleaner Production 202, 1145–1157.

Alcantara, V., Duro, J.A., 2004. Inequality of energy intensities across OECD countries: A note. Energy Policy 32, 1257–1260.

Ang, B.W., 2004. Decomposition analysis for policymaking in energy: Which is the preferred method? Energy Policy 32, 1131–1139.

Bianco, V., Cascetta, F., Marino, A., Nardini, S., 2019. Understanding energy consumption and carbon emissions in Europe: A focus on inequality issues. Energy 170,120-130.

Bourguignon, F., 1979. Decomposable income inequality measures. Econometrica 47, 901–920.

Bourguignon, F., Morrisson, C., 2002. Inequality among world citizens: 1820–1992. American Economic Review 92, 727-744.

Cowell, F. A., 2000. Measurement of inequality. In: Atkinson, A. B., Bourguignon F. (Eds.), Handbook of Income Distribution, Vol. I. North Holland, Amsterdam.

Cowell, F. A., 2011. Measuring inequality (3rd ed.). Oxford, England: Oxford University Press.

De La Peña, L., Guo, R., Cao, X., Ni, X., Zhang, W., 2022. Accelerating the energy transition to achieve carbon neutrality. Resources, Conservation and Recycling 177, 105957.

Duclos, J.-Y., Esteban, J., Ray, D., 2004. Polarization: Concepts, measurement, estimation. Econometrica 72, 1737–1772.

Duro, J.A., 2010. Decomposing international polarization of per capita CO2 emissions. Energy Policy 38, 6529-6533.

Duro, J.A., Padilla, E., 2006. International inequalities in per capita CO2 emissions: A decomposition methodology by Kaya factors. Energy Economics 28, 170-187.

Duro, J.A., Padilla, E., 2008. Analysis of the international distribution of per capita CO2 emissions using the polarization concept. Energy Policy 36, 456-466.

Duro, J.A., Padilla, E., 2011. Inequality across countries in energy intensities: an analysis of the role of energy transformation and final energy consumption. Energy Economics, 33, 474-479.

Duro, J.A., Teixidó-Figueras, J., Padilla, E., 2016. Empirics of the international inequality in CO2 emissions intensity: Explanatory factors according to complementary decomposition methodologies. Environmental and Resource Economics 63, 57-77.

Eskander, S.M.S.U., Nitschke, J., 2021. Energy use and CO2 emissions in the UK universities: An extended Kaya identity analysis. Journal of Cleaner Production 309, 127199.

Esteban, J.M., Ray, D., 1994. On the Measurement of Polarisation. Econometrica 62, 819-851.

Feng, K., Hubacek, K., Liu, Y., Marchan, E., Vogt-Schilb, A., 2018. Managing the distributional effects of energy taxes and subsidy removal in Latin America and the Caribbean. Applied Energy 225, 424-436.

Feng, K., Hubacek, K., Song, K., 2021. Household carbon inequality in the US. Journal of Cleaner Production 278, 123994.

Forbes, K.J., 2000. A reassessment of the relationship between inequality and growth. American Economic Review 90, 869–887.

Gottschalk, P., Smeeding, T.M., 1997. Cross-national comparisons of earnings and income inequality. Journal of Economic Literature 35, 633-687.

Greenstone, M., Hanna, R., 2014. Environmental regulations, air and water pollution, and infant mortality in India. American Economic Review 104, 3038-3072.

Groot, L.,2010. Carbon Lorenz curves. Resource and Energy Economics 32, 45-64.

Grunewald, N., Jacob, M., Mouratiadou, I., 2014. Decomposing inequality in CO2 emissions: The role of primary energy carriers and economic sectors. Ecological Economics 100, 183-194.

Hammond, G.P., Norman, J.B., 2012. Decomposition analysis of energy-related carbon emissions from UK manufacturing. Energy 41, 220-227.

Heil, M., Selden, T.M., 2001. Carbon emissions and economic development: Future trajectories based on historical experience. Environment and Development Economics 6, 63-83.

Heil, M.T., Wodon, Q.T., 1997. Inequality in CO2 emissions between poor and rich countries. Journal of Environment and Development 6, 426–452.

Heil, M.T., Wodon, Q.T., 2000. Future inequality in CO2 emissions and the impact of abatement proposals. Environmental and Resource Economics 17, 163-181.

Intergovernmental Panel on Climate Change (IPCC), 1997. Climate Change 1995. Economic and social dimensions of climate change. Contribution of WG III to the 2nd AR of the IPCC. Cambridge University Press, Cambridge.

Jorgenson, A., Schor, J., Huang, X., 2017. Income inequality and carbon emissions in the United States: A state-level analysis, 1997–2012. Ecological Economics 134, 40-48.

Kaya, Y., 1989. Impact of carbon dioxide emission control on GNP growth: interpretation of proposed scenarios. In: Paper Presented to the Energy and Industry Subgroup, Response Strategies Working Group. IPCC, Paris, France.

Li, H., Squire, L., Zou, H., 1998. Explaining international and intertemporal variations in income inequality. The Economic Journal 108, 26-43.

Lin, K.-H., Tomaskovic-Devey, D., 2013. Financialization and U.S. income inequality, 1970-2008. American Journal of Sociology 118, 1284-1329.

Madsen, J.B., Islam, M.R., Doucouliagos, H, 2018. Inequality, financial development and economic growth in the OECD, 1870–2011. European Economic Review 101, 605-624.

Mavromatidis, G., Orehounig, K., Richner, P., Carmeliet, J., 2016. A strategy for reducing CO2 emissions from buildings with the Kaya identity – A Swiss energy system analysis and a case study. Energy Policy 88, 343-354.

Padilla, E., Duro, J.A., 2013. Explanatory factors of CO2 per capita emission inequality in the European Union. Energy Policy 62, 1320-1328.

Padilla, E., Serrano, A., 2006. Inequality in CO2 emissions across countries and its relationship with income inequality: a distributive approach. Energy Policy 34, 1762–1772.

Patiño, L.I., Alcàntara, V., Padilla, E., 2021. Driving forces of CO2 emissions and energy intensity in Colombia. Energy Policy 151, 112130.

Ravallion, M. Heil, M., Jalan, J., 2000. Carbon emissions and income inequality. Oxford Economic Papers 52, 651-669.

Remuzgo, L., Sarabia, J.M., 2015. International inequality in CO2 emissions: A new factorial decomposition based on Kaya factors. Environmental Science and Policy, 54, 15-24.

Roux, N., Kastner, T., Erb, K.-H., Haberl, H., 2021. Does agricultural trade reduce pressure on land ecosystems? Decomposing drivers of the embodied human appropriation of net primary production. Ecological Economics 181, 106915.

Roux, N., Plank, B., 2022. The misinterpretation of structure effects of the LMDI and an alternative index decomposition. MethodsX 9, 101698.

Sarabia, J.M., Jorda, V., Remuzgo, L., 2017. The Theil indices in parametric families of income distributions - A short review. Review of Income and Wealth 63, 867-880.

Schmalensee, R., Stoker, T., Judson, R., 1998. World carbon dioxide emissions: 1950-2050. Review of Economics and Statistics 80, 15-27.

Shorrocks, A., 1980. The class of additively decomposable inequality measures. Econometrica 48, 613–625.

Shorrocks, A., 1984. Inequality decomposition by population subgroup. Econometrica 52, 1369-1385.

Shorrocks, A., Foster, J.E., 1987. Transfer sensitive inequality measures. Review of Economics Studies 54, 485–497.

Tavakoli, A., 2018. A journey among top ten emitter country, decomposition of "Kaya Identity". Sustainable Cities and Society 38, 254 - 264.

Theil, H., 1967. Economics and information theory. North-Holland, Amsterdam.

Theil, H., 1972. Statistical decomposition analysis with applications in the social and administrative sciences. North-Holland, Amsterdam.

Tian, Q., Zhao, T., Yuan, R., 2021. An overview of the inequality in China's carbon intensity 1997-2016: a Theil index decomposition analysis. Clean Technologies and Environmental Policy 23, 1581-1601.

Torras, M., Boyce, J., 1998. Income, inequality, and pollution: A reassessment of the environmental Kuznets curve. Ecological Economics 25, 147-160.

Ürge-Vorsatz, D., Cabeza, L.F., Serrano, S., Barreneche, C., Petrichenko, K., 2015. Heating and cooling energy trends and drivers in buildings. Renewable and Sustainable Energy Reviews 41, 85-98.

Wang, Y., Yan, Q., Li, Z., Baležentis, T., Zhang, Y., Gang, L., Štreimikienė, D., 2020. Aggregate carbon intensity of China's thermal electricity generation: The inequality analysis and nested spatial decomposition. Journal of Cleaner Production 247, 119139.

Wang, H., Zhou, P., 2018. Assessing global CO2 emission inequality from consumption perspective: An index decomposition analysis. Ecological Economics 154, 257-271.

Wolfson, M., 1994. When inequalities diverge. American Economic Review 94, 353-358.

York, R., Rosa, E.A., Dietz, T., 2003. A rift in modernity? Assessing the anthropogenic sources of global climate change with the STIRPAT model. International Journal of Sociology and Social Policy 23, 31-51.

Zaman, K., Ahmad, A., Hamzah T.A.A.T, Yusoff, M.M., 2016. Environmental factors affecting health indicators in sub-Saharan African countries: Health is wealth. Social Indicators Research 129, 215–228.

Zhang X., Kanbur R., 001. What difference do polarization measures make? An application to China. Journal of Development Studies 37, 85-98.

Zhu, Q., Peng, X., 2012. The impacts of population change on carbon emissions in China during 1978-2008. Environmental Impact Assessment Review 36, 1–8.

Table 1: Data & Sources

Panel A: Main variables

Variables	Unit	Time Series	Source
CO_2 Emissions	Million mt	1980-2017	US Energy Information Administration (eia.gov)
Fuel Consumption	Million Btu	1980-2017	US Energy Information Administration
Total Primary Energy Consumption	Million Btu	1980-2017	US Energy Information Administration
Gross Domestic Product	Million $	1980-2017	US Bureau of Labor Statistics (bls.gov)
Consumer Price Index[a]	Index	1980-2017	US Bureau of Labor Statistics
Employment	Thousand Persons	1980-2017	US Bureau of Labor Statistics
Population	Thousand Persons	1980-2017	US Bureau of Labor Statistics

Panel B: Other data

Variables	Time Series	Source
Geography[b]		
Broad regions	No	US Census Bureau (census.gov)
Divisions	No	Bureau of Economic Analysis
PADD districts	No	US Energy Information Administration
Geology[b]		
Elevation (mean)	No	US Geological Survey (USGS; usgs.gov)
Coastline (ocean)	No	National Oceanic and Atmospheric Admin. (NOAA)
Shoreline	No	NOAA (NOAA; shoreline.noaa.gov)
Forest cover	Irregular [c]	US Department of Agriculture (USDA; fs.usda.gov)
Climate[b]		
Temperature	1980-2017 [d]	NOAA (ncdc.noaa.gov)
Precipitation	1980-2017 [d]	NOAA
Drought (Palmer Severity Index)	1980-2017 [d]	NOAA
Heating Degree Days	1980-2017 [d]	NOAA
Cooing Degree Days	1980-2017 [d]	NOAA
Human Development[b]		
Education attainment [e]	2006-2017	US Census Bureau
Life expectancy at birth	1980-2017	US Mortality Database [f] (usa.mortality.org);
Real disposable income	1980-2017	Bureau of Economic Analysis
Development Indicator [g]	Based on the above	Authors' own calculations
Other [b]		
Energy cons. vs prod.	1980-2017	US Energy Information Administration
Fossil fuel prod. vs. cons	1980-2017	US Energy Information Administration
Manufacturing GDP (% of Total GDP)	1980-2017	Bureau of Economic Analysis
Population density	1980-2017	Bureau of Economic Analysis

Notes:

[a] CPI is obtained for four broad regions: Northeastern, South, West, Midwest; this is used to convert regional (state) nominal gross domestic products and real disposable incomes to real.

[b] Groups that are based on geography and geology are static and data depend on the measurement period. For example, coastline figures were measured in 1915 and measured again in 1948, 1961 and 1975; with only few changes; these data are not expected to be different within the course of this study's timeframe. On the other hand, forest cover, might have been different throughout the years; for instance, the state of Mississippi had recorded 16.7 million acres of forest land in 1977 and 19.4 in 2017, representing an increase of more than 16%; for this variable we calculate the average per state forest cover (forest/total area) from 1977 to 2017. The same procedure is for the climate and human development indicators, i.e., groups are based on US state annual averages throughout time.

[c] Data on forest cover are available for the years 1977, 1987, 1997, 2007, 2012 and 2017.

[d] Data are obtained for monthly frequencies and are converted to annual averages.

[e] Education attainment refers to % of the total the population (aged 18 years or above) with a Bachelor's degree or higher and are not available prior to 2006; these data come from the ACS (American Community Surveys) 1- and 5- year Estimates.

[f] University of California, Berkeley.

[g] Human Development Indicator is constructed as the product of three standardized variables: education attainment, life expectancy at birth and real disposable income; this follows the United Nations Development Program which identified health, education, and material wellbeing as the key factors of human development, and further combined these three factors into a single measure, the human development index (HDI; UNDP, 2013. Human Development Report 2013. New York: United Nations).

Table 2. Factorial decomposition of inequality in CO_2 emissions

Year	T	T^a	T^b	T^c	T^d	T^e	$\ln\left(\frac{\bar{y}}{y_i^a}\right)$	$\ln\left(\frac{\bar{y}}{y_i^b}\right)$	$\ln\left(\frac{\bar{y}}{y_i^c}\right)$	$\ln\left(\frac{\bar{y}}{y_i^d}\right)$
1980	7.50	0.29 (3.93)	0.87 (11.6)	5.57 (74.3)	0.76 (10.1)	0.16 (2.13)	-1.07 (-14.2)	1.20 (16.0)	-0.26 (-3.52)	-0.03 (-0.35)
1981	8.01	0.28 (3.50)	0.99 (12.4)	5.68 (70.9)	0.90 (11.2)	0.17 (2.06)	-0.90 (-11.3)	1.11 (13.8)	-0.20 (-2.46)	-0.01 (-0.13)
1982	7.72	0.28 (3.65)	1.10 (14.3)	5.97 (77.3)	0.89 (11.6)	0.19 (2.42)	-0.72 (-9.34)	1.01 (13.0)	-0.98 (-12.6)	-0.02 (-0.27)
1983	8.08	0.29 (3.56)	1.26 (15.6)	6.74 (83.5)	0.73 (8.99)	0.20 (2.51)	-0.45 (-5.57)	1.10 (13.6)	-1.71 (-21.2)	-0.08 (-1.02)
1984	8.12	0.31 (3.82)	1.15 (14.2)	7.18 (88.4)	0.68 (8.41)	0.19 (2.35)	-0.47 (-5.76)	1.10 (13.6)	-1.96 (-24.2)	-0.07 (-0.85)
1985	8.03	0.33 (4.14)	1.01 (12.6)	7.20 (89.6)	0.67 (8.30)	0.17 (2.09)	-0.29 (-3.57)	1.23 (15.3)	-2.26 (-28.2)	-0.02 (-0.23)
1986	7.99	0.31 (3.82)	1.05 (13.1)	8.44 (105)	0.66 (8.24)	0.16 (2.05)	-0.23 (-2.90)	1.08 (13.5)	-3.50 (-43.7)	0.02 (0.31)
1987	7.78	0.32 (4.15)	0.87 (11.2)	8.75 (112)	0.66 (8.47)	0.16 (2.08)	-0.30 (-3.89)	1.12 (14.4)	-3.86 (-49.6)	0.06 (0.71)
1988	8.34	0.30 (3.63)	0.83 (9.95)	9.21 (110)	0.68 (8.11)	0.15 (1.81)	-0.22 (-2.59)	1.22 (14.6)	-3.86 (-46.3)	0.03 (0.36)
1989	8.37	0.31 (3.71)	0.83 (9.94)	9.04 (108)	0.64 (7.62)	0.13 (1.59)	-0.28 (-3.34)	1.23 (14.7)	-3.57 (-42.7)	0.04 (0.47)
1990	9.13	0.34 (3.68)	0.93 (10.2)	9.32 (102)	0.63 (6.86)	0.11 (1.22)	-0.23 (-2.50)	1.25 (13.7)	-3.29 (-36.1)	0.07 (0.75)
1991	9.11	0.35 (3.87)	0.87 (9.54)	9.05 (99.4)	0.58 (6.33)	0.12 (1.35)	-0.22 (-2.38)	1.28 (14.1)	-2.93 (-32.2)	0.00 (-0.02)
1992	9.23	0.36 (3.95)	0.72 (7.84)	9.03 (97.9)	0.52 (5.61)	0.14 (1.47)	-0.11 (-1.20)	1.44 (15.5)	-2.82 (-30.6)	-0.05 (-0.51)
1993	9.73	0.36 (3.69)	0.77 (7.90)	8.83 (90.8)	0.51 (5.28)	0.15 (1.56)	0.04 (0.38)	1.68 (17.3)	-2.50 (-25.7)	-0.11 (-1.15)
1994	9.43	0.36 (3.78)	0.70 (7.45)	8.73 (92.5)	0.50 (5.35)	0.17 (1.78)	-0.03 (-0.35)	1.46 (15.5)	-2.32 (-24.6)	-0.14 (-1.49)
1995	9.63	0.36 (3.70)	0.81 (8.46)	8.73 (90.7)	0.54 (5.62)	0.18 (1.85)	0.13 (1.38)	1.41 (14.6)	-2.35 (-24.4)	-0.19 (-1.96)
1996	9.93	0.35 (3.56)	0.88 (8.88)	9.20 (92.6)	0.55 (5.57)	0.17 (1.67)	0.05 (0.47)	1.34 (13.5)	-2.46 (-24.7)	-0.16 (-1.61)
1997	10.0	0.38 (3.77)	0.89 (8.94)	10.1 (100)	0.67 (6.69)	0.15 (1.49)	0.03 (0.30)	1.20 (12.0)	-3.25 (-32.4)	-0.11 (-1.13)
1998	9.85	0.40 (4.05)	0.70 (7.12)	10.4 (106)	0.76 (7.67)	0.14 (1.41)	0.19 (1.90)	1.43 (14.5)	-4.13 (-41.9)	-0.07 (-0.70)
1999	9.61	0.41 (4.22)	0.77 (8.04)	10.5 (110)	0.83 (8.69)	0.13 (1.35)	0.26 (2.70)	1.28 (13.3)	-4.55 (-47.4)	-0.05 (-0.56)
2000	9.70	0.38 (3.96)	0.66 (6.84)	11.0 (113)	0.94 (9.74)	0.11 (1.18)	0.21 (2.13)	1.40 (14.4)	-5.01 (-51.6)	-0.03 (-0.27)
2000	9.70	0.38 (3.96)	0.66 (6.84)	11.0 (113)	0.94 (9.74)	0.11 (1.18)	0.21 (2.13)	1.40 (14.4)	-5.01 (-51.6)	-0.03 (-0.27)
2001	9.27	0.36 (3.92)	0.56 (6.05)	10.6 (114)	0.95 (10.2)	0.12 (1.24)	0.27 (2.87)	1.33 (14.4)	-4.87 (-52.5)	-0.03 (-0.28)
2002	9.85	0.36 (3.64)	0.72 (7.28)	11.1 (113)	0.88 (8.98)	0.12 (1.25)	0.32 (3.26)	1.38 (14.0)	-5.00 (-50.8)	-0.02 (-0.23)
2003	9.72	0.35 (3.58)	0.68 (6.99)	10.4 (107)	0.84 (8.68)	0.13 (1.29)	0.47 (4.82)	1.41 (14.5)	-4.51 (-46.4)	-0.06 (-0.61)
2004	9.31	0.35 (3.73)	0.62 (6.63)	9.96 (107)	0.81 (8.71)	0.11 (1.23)	0.53 (5.72)	1.17 (12.6)	-4.21 (-45.3)	-0.03 (-0.36)
2005	9.12	0.34 (3.70)	0.60 (6.59)	9.30 (102)	0.80 (8.80)	0.11 (1.22)	0.73 (7.99)	1.19 (13.1)	-3.92 (-43.0)	-0.03 (-0.31)
2006	9.74	0.35 (3.61)	0.68 (7.03)	9.53 (97.8)	0.86 (8.79)	0.11 (1.12)	0.76 (7.76)	1.34 (13.7)	-3.85 (-39.5)	-0.03 (-0.36)
2007	9.43	0.35 (3.74)	0.62 (6.55)	9.39 (99.6)	0.88 (9.35)	0.12 (1.23)	0.74 (7.83)	1.12 (11.9)	-3.78 (-40.1)	-0.01 (-0.10)
2008	9.83	0.35 (3.57)	0.65 (6.61)	9.15 (93.1)	0.85 (8.65)	0.13 (1.31)	0.96 (9.78)	1.12 (11.4)	-3.41 (-34.7)	0.03 (0.33)
2009	9.54	0.33 (3.51)	0.69 (7.23)	9.46 (99.2)	0.87 (9.13)	0.17 (1.75)	0.89 (9.34)	1.01 (10.6)	-4.02 (-42.1)	0.13 (1.33)
2010	10.2	0.36 (3.50)	0.67 (6.56)	9.89 (96.7)	0.92 (8.96)	0.16 (1.59)	0.92 (8.96)	1.08 (10.6)	-3.89 (-38.0)	0.12 (1.20)
2011	11.1	0.36 (3.20)	0.86 (7.73)	10.0 (90.0)	0.96 (8.62)	0.16 (1.43)	1.02 (9.14)	1.28 (11.5)	-3.61 (-32.5)	0.10 (0.88)
2012	10.9	0.34 (3.16)	0.87 (8.01)	10.3 (94.7)	1.10 (10.1)	0.15 (1.35)	0.90 (8.28)	1.01 (9.24)	-3.91 (-35.8)	0.12 (1.10)
2013	10.9	0.34 (3.15)	0.81 (7.38)	10.4 (94.8)	1.09 (9.96)	0.15 (1.33)	0.91 (8.28)	1.08 (9.86)	-3.95 (-36.2)	0.15 (1.37)
2014	10.8	0.33 (3.04)	0.80 (7.42)	10.6 (97.9)	1.18 (10.9)	0.15 (1.38)	0.93 (8.58)	0.96 (8.93)	-4.27 (-39.6)	0.15 (1.40)
2015	10.0	0.30 (2.96)	0.74 (7.32)	11.1 (110)	1.20 (12.0)	0.14 (1.43)	0.63 (6.25)	0.94 (9.33)	-5.14 (-51.1)	0.18 (1.82)
2016	9.96	0.26 (2.64)	0.74 (7.41)	11.8 (118)	1.29 (13.0)	0.14 (1.38)	0.51 (5.17)	0.96 (9.67)	-5.91 (-59.3)	0.17 (1.74)
2017	10.4	0.25 (2.44)	0.76 (7.32)	11.9 (115)	1.39 (13.5)	0.13 (1.26)	0.58 (5.56)	1.10 (10.6)	-5.88 (-56.8)	0.13 (1.30)
Avg	9.35	0.34 (3.61)	0.81 (8.85)	9.30 (99.1)	0.82 (8.76)	0.15 (1.59)	0.20 (1.58)	1.21 (13.1)	-3.47 (-36.6)	0.00 (0.02)
5% $_{tail}$	7.77	0.28 (2.91)	0.62 (6.56)	5.92 (76.9)	0.52 (5.54)	0.11 (1.21)	-0.75 (-9.63)	0.96 (9.32)	-5.25 (-53.2)	-0.14 (-1.50)
95% $_{tail}$	10.9	0.39 (4.14)	1.11 (14.2)	11.2 (114)	1.22 (12.1)	0.19 (2.36)	0.93 (9.17)	1.44 (15.6)	-0.87 (-11.3)	0.15 (1.45)

All figures are multiplied by 100 for exposition purposes. T is the Theil index. T^f is the partial contribution of each factor to total inequality T, for $f = \{a = CO_2/FC; b = FC/EC; c = EC/GDP; d = GDP/EP; e = EP/TP\}$. CO_2, *FC*, *EC*, *GDP*, *EP* and *TP* are, respectively, the carbon emissions, fuel consumption, total primary energy consumption, real gross domestic (state) product, employment and total population. Avg, 5% and 95% tails are the average, 5% and 95% percentiles calculated across the period 1980 to 2017 using 38 annual observations. Figures in parentheses show the contribution to Theil index in percentage terms.

Table 3: Within and between group results for groups based on location and CO_2 per capita distribution

	T_j/T	T^a/T	T^b/T	T^c/T	T^d/T	T^e/T	$\ln\left(\frac{\bar{y}}{y_i^a}\right)$	$\ln\left(\frac{\bar{y}}{y_i^b}\right)$	$\ln\left(\frac{\bar{y}}{y_i^c}\right)$	$\ln\left(\frac{\bar{y}}{y_i^d}\right)$	$ZK = T_B/T_W$	
Panel A: Geographical location												
Census Regions (four-polar)												
T_B	28.17	1.96	4.02	111.51	6.11	1.01	3.80	24.19	-54.14	1.54	$\overline{ZK}$	0.39
	{-1}	{1}	{-8}	{100}	{46}	{1}	{75}	{-2}	{-220}	{7}	90%CI	0.34/0.45
T_W	71.83	4.07	11.41	98.60	9.68	1.81	0.86	8.05	-34.02	-0.45	r	[-0.06]
	{1}	{-6}	{-14}	{111}	{-5}	{-4}	{46}	{-18}	{-115}	{4}		
Census Divisions (nine-polar)												
T_B	60.03	4.80	4.54	94.42	6.39	1.20	-10.21	30.24	-31.22	-0.14	$\overline{ZK}$	1.51
	{13}	{-8}	{-7}	{109}	{9}	{-1}	{94}	{-19}	{-180}	{3}	90%CI	1.25/1.68
T_W	39.97	2.91	17.67	115.20	12.03	2.15	17.98	-15.03	-53.48	0.57	r	[0.56]
	{-13}	{0}	{-18}	{111}	{9}	{-4}	{6}	{-18}	{-94}	{9}		
PADD Districts (five-polar)												
T_B	54.57	3.48	4.44	92.67	2.21	0.60	-12.42	30.81	-20.97	-0.81	$\overline{ZK}$	1.22
	{32}	{-7}	{-6}	{86}	{1}	{0}	{95}	{-5}	{-164}	{0}	90%CI	0.91/1.53
T_W	45.43	4.76	15.99	117.74	16.79	2.72	16.94	-10.20	-65.91	1.15	r	[1.28]
	{-32}	{-2}	{-12}	{166}	{32}	{-4}	{27}	{-49}	{-170}	{13}		
Panel B: CO_2 per capita distribution												
I. Bi-polar: above/below median												
T_B	62.89	0.16	3.56	86.59	1.67	0.16	6.83	29.49	-29.17	0.71	$\overline{ZK}$	1.71
	{-9}	{0}	{-3}	{80}	{8}	{0}	{28}	{-15}	{-104}	{5}	90%CI	1.45/1.98
T_W	37.11	9.09	20.07	127.79	20.52	4.03	-7.26	-17.07	-56.00	-1.17	r	[-0.36]
	{9}	{-13}	{-30}	{156}	{4}	{-6}	{97}	{-3}	{-209}	{5}		
II. Other Bi-polar - 26·22 -												
T_B	63.27	0.06	3.88	85.62	1.45	0.14	3.03	31.89	-26.65	0.58	$\overline{ZK}$	1.73
	{5}	{0}	{-2}	{78}	{5}	{0}	{34}	{-14}	{-104}	{4}	90%CI	1.56/1.91
T_W	36.73	9.47	19.16	129.91	21.26	4.06	-0.68	-21.15	-61.19	-0.84	r	[0.23]
	{-5}	{-11}	{-27}	{178}	{20}	{-5}	{88}	{-22}	{-227}	{5}		
III. Tripolar: Tertiles												
T_B	80.74	0.85	3.21	81.47	1.92	0.13	4.75	28.48	-20.83	0.03	$\overline{ZK}$	4.23
	{3}	{0}	{-4}	{102}	{7}	{0}	{41}	{-21}	{-128}	{2}	90%CI	3.55/4.88
T_W	19.26	15.09	35.29	188.24	37.03	7.64	-11.39	-54.52	-117.5	0.13	r	[0.16]
	{-3}	{-20}	{-45}	{183}	{21}	{-10}	{99}	{9}	{-251}	{14}		
IV. Other Tripolar - 14·22·12 -												
T_B	84.06	1.38	2.51	86.58	2.24	0.20	2.55	26.92	-22.31	-0.07	$\overline{ZK}$	5.38
	{-7}	{-1}	{-3}	{109}	{9}	{0}	{46}	{-16}	{-146}	{3}	90%CI	4.26/6.54
T_W	15.94	16.07	46.30	183.77	42.83	8.94	-4.13	-64.53	-129.7	0.44	r	[-0.48]
	{7}	{-28}	{-73}	{123}	{-10}	{-17}	{97}	{31}	{-137}	{12}		
V. Quadripolar: Quantiles												
T_B	86.81	1.19	2.84	86.09	2.20	0.27	2.73	27.39	-22.80	0.08	$\overline{ZK}$	6.79
	{-8}	{-1}	{-3}	{102}	{7}	{0}	{42}	{-16}	{-134}	{2}	90%CI	4.99/8.40
T_W	13.19	20.60	54.66	207.88	51.69	10.47	-7.06	-89.16	-148.2	-0.91	r	[-0.55]
	{8}	{-31}	{-83}	{156}	{-6}	{-18}	{120}	{42}	{-197}	{17}		
VI. Quadripolar: Quantiles -10·10·16·12-												
T_B	87.31	1.58	3.23	82.86	3.33	0.23	3.01	28.95	-22.92	-0.27	$\overline{ZK}$	7.05
	{-6}	{-1}	{-4}	{99}	{14}	{0}	{45}	{-10}	{-147}	{3}	90%CI	5.38/8.65
T_W	12.69	18.74	55.15	231.09	45.99	11.09	-9.49	-105.29	-149.5	2.21	r	[-0.48]
	{6}	{-32}	{-82}	{140}	{-34}	{-19}	{109}	{12}	{-108}	{14}		
VII. Other multipolar: -8·8·12·12·8-												
T_B	89.75	1.66	3.48	79.05	3.41	0.25	3.78	28.42	-19.48	-0.57	$\overline{ZK}$	9.20
	{-10}	{-1}	{-3}	{98}	{11}	{0}	{40}	{-13}	{-132}	{0}	90%CI	6.02/12.6
T_W	10.25	21.92	65.96	302.46	56.25	13.73	-19.41	-135.44	-210.4	4.95	r	[-0.88]
	{10}	{-46}	{-126}	{94}	{-41}	{-28}	{175}	{85}	{-144}	{31}		

Numbers in {} are annual absolute changes in basis points (bps). Figures in [] represent % growth rates per annum.

Table 4. Bipolar groups: Between-group inequality in CO_2 emissions and explanatory power

	Partition	T_W/T	T_B/T	*(90% CI)*	ΔT_B	$\overline{ZK}$	*(90% CI)*	*r*
Kaya Factor I: CO_2/FC (a)	38·10	83.84	**16.16**	(9.95 20.1)	{13}	**0.19**	(0.11 0.25)	[1.34]
Kaya Factor II: FC/EC (b)	34·14	49.17	**50.83*****	(44.9 57.3)	{-25}	**1.05*****	(0.82 1.34)	[-1.01]
Kaya Factor III: EC/GDP (c)	36·12	42.69	**57.31*****	(51.6 62.5)	{-2.3}	**1.35*****	(1.07 1.67)	[-0.09]
Kaya Factor IV: GDP/EP (d)	38·10	91.53	**8.47**	(4.61 12.4)	{4.0}	**0.09**	(0.05 0.14)	[0.96]
Kaya Factor V: EP/TP (e)	10·38	85.28	**14.72**	(12.3 17.6)	{9.2}	**0.17**	(0.14 0.21)	[0.78]
Other I: Energy prod. vs. cons.	34·14	52.39	**47.61*****	(42.8 53.1)	{-19}	**0.92*****	(0.75 1.13)	[-0.76]
Other II: Fossil fuel prod. vs. cons.	36·12	53.62	**46.38*****	(40.0 52.3)	{-14}	**0.87*****	(0.67 1.09)	[-0.58]
Other III: Manufact. GDP (% of GDP)	16·32	82.26	**17.74***	(15.5 21.0)	{1.2}	**0.22***	(0.18 0.27)	[0.09]
Other IV: Population density	34·14	63.35	**36.65*****	(30.6 40.3)	{28}	**0.58*****	(0.44 0.67)	[1.29]
Geology I: Mean elevation	38·10	92.13	**7.87**$^{+}$	(4.96 10.6)	{7.6}	**0.09**	(0.05 0.12)	[1.31]
Geology II: Forest cover	16·32	75.68	**24.32****	(21.2 27.0)	{-5.8}	**0.32****	(0.27 0.37)	[-0.31]
Geology III: Coastline/shoreline	30·18	94.54	5.46	(1.54 9.80)	{13}	0.06	(0.02 0.11)	[4.29]
Climate I: Temperature (average)	38·10	88.96	**11.04**	(6.99 16.2)	{-18}	**0.13**	(0.08 0.19)	[-1.59]
Climate II: Precipitation index	10·38	96.09	3.91	(2.31 5.91)	{3.0}	0.04	(0.02 0.06)	[1.07]
Climate III: Drought severity (Palmer)	14,34	84.08	**15.92**	(10.6 22.6)	{34}	**0.19**	(0.12 0.29)	[2.52]
Climate IV: Heating degree days	28·20	95.64	4.36	(0.67 11.2)	{-30}	0.05	(0.01 0.13)	[-7.68]
Climate V: Cooling degree days	26·22	83.39	**16.61***	(10.7 20.2)	{-21}	**0.20***	(0.12 0.25)	[-1.73]
Human develop. I: Educ. attain.	34·14	55.64	**44.36*****	(38.8 47.7)	{-8.9}	**0.80*****	(0.63 0.91)	[-0.37]
Human develop. II: Life exp.	24·24	62.22	**37.78*****	(30.8 43.9)	{-38}	**0.61*****	(0.44 0.78)	[-1.62]
Human develop.III: Real disp. income	36·12	62.70	**37.30*****	(32.4 40.3)	{-7.6}	**0.60*****	(0.48 0.68)	[-0.34]
Human develop.IV: HDI indicator	36·12	57.08	**42.92*****	(37.7 46.2)	{-11}	**0.76*****	(0.61 0.86)	[-0.48]

T is the Theil index. T^f is the partial contribution of each factor to total inequality T, for $f = \{a = CO_2/FC; b = FC/EC; c = EC/GDP; d = GDP/EP; e = EP/TP\}$. *$CO_2$*, *FC*, *EC*, *GDP*, *EP* and *TP* are, respectively, the carbon emissions, fuel consumption, total primary energy consumption, real gross domestic (state) product, employment and total population. Avg, 5% and 95% tails are the average, 5% and 95% percentiles calculated across the period 1980 to 2017 using 38 annual observations. Figures in parentheses show the contribution to Theil index in percentage terms. Asterisks ***, **, * indicate that the corresponding explanatory capacity (T_B/T) of the defined grouping exceeds the right-side of the 99, 95, 90% confidence interval of T_B/T; the distribution of this is approximated by randomly assigning states into bipolar groups. The values are 16.50, 20.62, 28.57%. This is calculated using 3000 random permutations of the 48 contiguous states which are further allocated to 15 partitions resulting in 45000 (e.g., equal to 15 x 3000) random bipolar groups. Numbers in bold indicate whether the corresponding explanatory capacity (T_B/T) of the defined grouping is within the upper quartile of the distribution of the simulated T_B/T. Similar is the interpretation of asterisks attached to $\overline{ZK}$; the right-side of the 99, 95, 90% confidence intervals from the simulations are 0.198, 0.261 and 0.403.

Table 5. Tripolar groups: Between-group inequality in CO_2 emissions and explanatory power

	Partition	T_W/T	T_B/T	*(90% CI)*	ΔT_B	$\overline{ZK}$	*(90% CI)*	*r*
Kaya Factor I: CO_2/FC (a)	28·10·10	81.73	**18.27**	(11.9 23.3)	{8.3}	**0.23**	(0.14 0.30)	[0.77]
Kaya Factor II: FC/EC (b)	22·10·16	47.45	**52.55*****	(47.2 57.4)	{-20}	**1.12*****	(0.89 1.35)	[-0.79]
Kaya Factor III: EC/GDP (c)	10·26·12	21.68	**78.32*****	(75.6 81.6)	{8.1}	**3.66*****	(3.10 4.43)	[0.44]
Kaya Factor IV: GDP/EP (d)	12·26·10	90.24	9.76	(5.02 14.0)	{7.4}	0.11	(0.05 0.16)	[1.53]
Kaya Factor V: EP/TP (e)	10·12·26	84.48	**15.52**	(12.4 19.6)	{9.2}	**0.18**	(0.14 0.24)	[0.73]
Other I: Energy prod. vs. cons.	24·10·14	48.08	**51.92*****	(46.4 57.5)	{-19}	**1.09*****	(0.87 1.35)	[-0.74]
Other II: Fossil fuel prod. vs. cons.	22·14·12	50.03	**49.97*****	(42.9 57.9)	{-30}	**1.02*****	(0.75 1.38)	[-1.21]
Other III: Manufact. GDP (% of GDP)	16·14·18	81.71	**18.29*****	(15.9 21.2)	{2.6}	**0.22**	(0.19 0.27)	[0.18]
Other IV: Population density	20·14·14	61.72	**38.28*****	(34.4 42.2)	{13}	**0.62*****	(0.52 0.73)	[0.57]
Geology I: Mean elevation	28·10·10	80.88	**19.12**	(16.2 21.8)	{2.7}	**0.24**	(0.19 0.28)	[0.17]
Geology II: Forest cover	16·10·22	69.19	**30.81****	(27.8 33.7)	{-8.6}	**0.45****	(0.39 0.51)	[-0.42]
Geology III: Coastline/shoreline	20·10·18	94.41	5.59	(2.30 9.87)	{9.1}	0.06	(0.02 0.11)	[2.27]
Climate I: Temperature (average)	26·12·10	85.83	**14.17**	(10.5 17.4)	{-8.9}	**0.17**	(0.12 0.21)	[-0.67]
Climate II: Precipitation index	10·10·28	91.16	8.84	(6.92 11.5)	{10}	0.10	(0.07 0.13)	[1.29]
Climate III: Drought severity (Palmer)	14·14·20	77.35	**22.65**	(19.9 27.3)	{15}	**0.29**	(0.25 0.38)	[0.79]
Climate IV: Heating degree days	28·10·10	91.05	8.95	(5.57 13.7)	{-21}	0.10	(0.06 0.16)	[-2.41]
Climate V: Cooling degree days	16·10·22	81.84	**18.16**	(12.6 21.4)	{-16}	**0.22**	(0.14 0.27)	[-1.24]
Human develop. I: Educ. attain.	12·22·14	47.53	**52.47*****	(47.2 55.9)	{0.7}	**1.11*****	(0.89 1.27)	[0.03]
Human develop. II: Life exp.	24·12·12	55.84	**44.16*****	(39.1 48.2)	{-17}	**0.79*****	(0.64 0.93)	[-0.70]
Human develop.III: Real disp. income	24·12·12	57.63	**42.37*****	(37.0 45.7)	{-7.7}	**0.74*****	(0.59 0.84)	[-0.33]
Human develop.IV: HDI indicator	12·24·12	51.46	**48.54*****	(43.5 52.6)	{-4.7}	**0.95*****	(0.77 1.11)	[-0.19]

See notes in Table 3. Asterisks ***, **, * indicate that the corresponding explanatory capacity (T_B/T) of the defined grouping exceeds the right-side of the 99, 95, 90% confidence interval of T_B/T; the distribution of this is approximated by randomly assigning states into tripolar and quadripolar groups. The values are 23.64, 27.48, 34.77% and 28.78, 32.70, 39.27 for tripolar and quadripolar groups, respectively. This is calculated using 3000 random permutations of the 48 contiguous states which are further allocated to 55 partitions resulting in 165000 random tripolar (= 55 x 3000) groups. Numbers in bold indicate whether the corresponding explanatory capacity (T_B/T) of the defined grouping is within the upper quartile of the distribution of the simulated T_B/T. Similar is the interpretation of asterisks attached to $\overline{ZK}$; the right-side of the 99, 95, 90% confidence intervals from the simulations are 0.311, 0.381 and 0.535. Note that we repeat the above experiments using quadpolar groups. This results in 105000 groups (= 35 x 3000) but results are qualitatively similar, hence not presented here for brevity.

Table 6: Further within and between group results

	T_j/T	T^a/T	T^b/T	T^c/T	T^d/T	T^e/T	$\ln\left(\frac{\bar{y}}{y_i^a}\right)$	$\ln\left(\frac{\bar{y}}{y_i^b}\right)$	$\ln\left(\frac{\bar{y}}{y_i^c}\right)$	$\ln\left(\frac{\bar{y}}{y_i^d}\right)$
Panel A: Kaya factors										
I. CO_2/FC (tripolar)										
T_B	18.27	15.03	1.92	77.19	10.28	0.32	42.45	9.55	-58.63	1.89
	{8}	{-42}	{5}	{93}	{6}	{-3}	{-2}	{1}	{-43}	{-15}
T_W	81.73	1.93	10.43	105.67	8.42	1.87	-8.16	13.76	-33.71	-0.22
	{-8}	{-2}	{-12}	{128}	{9}	{-2}	{58}	{-17}	{-168}	{7}
II. FC/EC (bipolar)										
T_B	50.83	0.16	6.20	68.61	0.51	0.08	-8.40	42.56	-9.86	0.14
	{-25}	{-1}	{6}	{59}	{-4}	{0}	{41}	{6}	{-111}	{3}
T_W	49.17	7.40	12.73	137.66	17.24	3.19	11.49	-18.38	-71.17	-0.15
	{25}	{-12}	{-36}	{129}	{12}	{-7}	{64}	{-7}	{-147}	{5}
III. EC/GDP (tripolar)										
T_B	78.32	1.07	1.96	104.14	3.00	0.31	0.79	19.13	-30.29	-0.11
	{8}	{0}	{3}	{93}	{10}	{0}	{50}	{-9}	{-150}	{4}
T_W	21.68	13.36	34.58	87.88	29.23	6.20	3.98	-9.68	-66.09	0.54
	{-8}	{-15}	{-47}	{179}	{16}	{-8}	{68}	{-39}	{-162}	{8}
IV. GDP/EP (quadripolar)										
T_B	12.24	17.23	3.47	193.09	53.50	2.58	40.06	-8.74	-207.1	5.91
	{-3}	{-10}	{-20}	{494}	{269}	{-8}	{17}	{-41}	{-648}	{-52}
T_W	87.76	1.93	9.57	87.66	2.82	1.47	-3.95	16.16	-15.04	-0.62
	{3}	{-4}	{-11}	{94}	{-11}	{-2}	{59}	{-13}	{-123}	{10}
V. EP/TP (bipolar)										
T_B	14.72	0.28	0.10	207.17	8.26	3.00	9.50	4.46	-141.0	8.28
	{9}	{0}	{0}	{215}	{43}	{-5}	{50}	{-4}	{-327}	{28}
T_W	85.28	4.08	10.33	83.48	8.81	1.36	0.34	14.51	-21.52	-1.39
	{-9}	{-5}	{-12}	{87}	{4}	{-2}	{54}	{-16}	{-110}	{0}
Panel B: Other partitioning										
I.. Energy prod. vs. cons. (quadripolar)										
T_B	59.93	0.88	5.11	91.10	1.68	0.23	-2.01	27.77	-25.00	0.25
	{-14}	{-3}	{0}	{75}	{4}	{0}	{36}	{-8}	{-105}	{1}
T_W	40.07	7.91	16.98	117.03	19.13	3.62	6.49	-11.46	-59.35	-0.36
	{14}	{-10}	{-35}	{158}	{7}	{-7}	{81}	{-12}	{-191}	{9}
II. Geology: Forest cover (tripolar)										
T_B	30.81	1.27	4.01	82.03	1.29	0.60	1.66	24.55	-16.28	0.87
	{-9}	{-5}	{-6}	{152}	{-2}	{-5}	{48}	{-45}	{-135}	{-2}
T_W	69.19	4.51	11.30	108.02	12.01	2.02	1.62	7.59	-46.76	-0.33
	{9}	{-4}	{-16}	{97}	{12}	{-1}	{56}	{1}	{-153}	{8}
III. Climate: Drought severity (quadripolar)										
T_B	29.30	1.72	2.41	95.56	4.00	1.25	-11.05	22.53	-14.08	-2.34
	{31}	{0}	{-1}	{-32}	{1}	{0}	{56}	{5}	{-32}	{4}
T_W	70.70	4.70	11.90	101.73	10.78	1.70	6.45	8.67	-46.97	1.03
	{-31}	{-6}	{-11}	{169}	{18}	{-3}	{59}	{-28}	{-204}	{6}
IV. Human development: Education attainment (tripolar)										
T_B	52.47	0.92	1.74	121.54	7.31	0.64	16.90	18.43	-70.16	2.67
	{1}	{1}	{-1}	{144}	{44}	{-2}	{22}	{-6}	{-214}	{11}
T_W	47.53	5.99	17.75	78.04	10.04	2.66	-14.98	6.30	-3.18	-2.62
	{-1}	{-9}	{-21}	{77}	{-23}	{-3}	{81}	{-22}	{-81}	{0}

See Table 3 for more details on groups. Group partitioning is reported on the basis of whether further partitioning, i.e., bipolar vs tripolar (tripolar vs quadripolar) offers an improvement in explanatory power of more than 10% (20%).

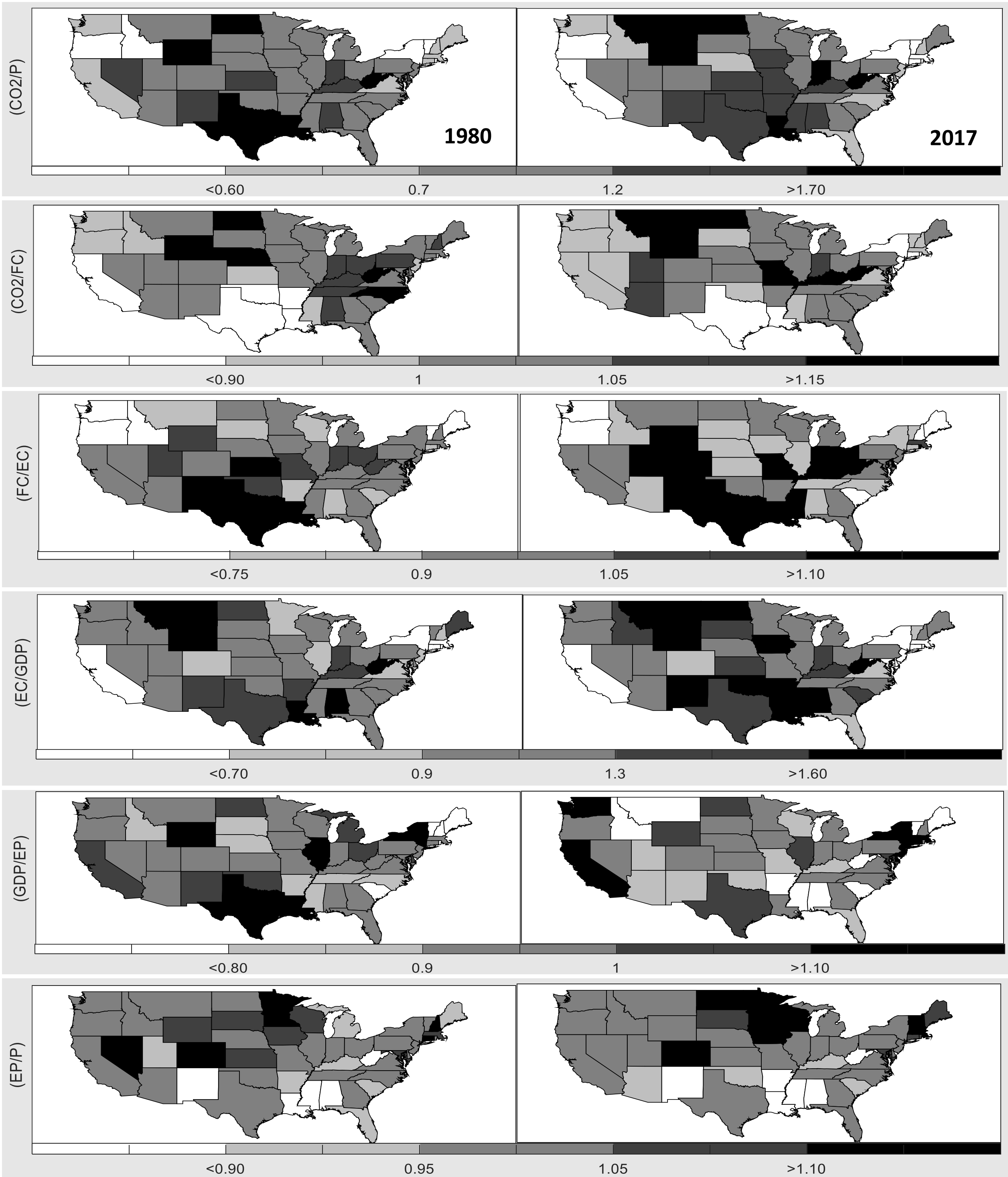

Figure 1: Differences CO_2 per capita and Kaya factor intensities across the US. Each variable Y = {CO_2/P, CO_2/FC, FC/EC, EC/GDP, GDP/P, EP/P}is calculated as Y_{STATE}/Y_{US}; Y_{US} is the population weighted US average. States are classified based on pre-defined thresholds derived from the 1980 percentiles of Y_{STATE}/Y_{US}, i.e., 10% (white) and 90% (black); 25% (light grey) and 75% (dark grey).

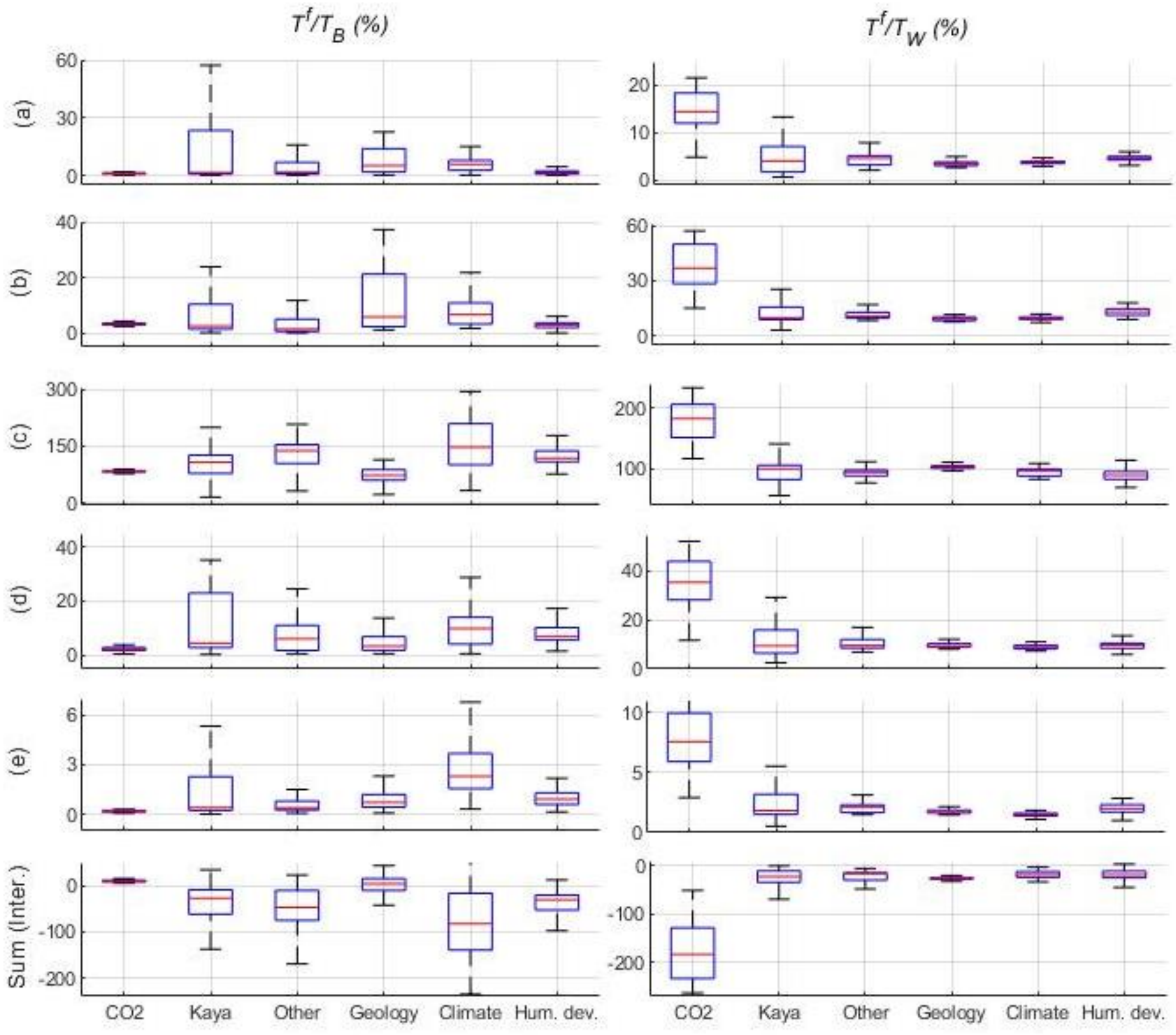

Figure 2: Factor contribution to between and within group inequality.