

Article

Evaluating the Causal Relations between the Kaya Identity Index and ODIAC-Based Fossil Fuel CO₂ Flux

YoungSeok Hwang ¹, Jung-Sup Um ² , JunHwa Hwang ¹  and Stephan Schlüter ^{3,*} 

¹ Department of Climate Change, Kyungpook National University, 80 University Road, Buk-gu, Daegu 41566, Korea; poiu01234@knu.ac.kr (Y.H.); jhhwang1907@knu.ac.kr (J.H.)

² Department of Geography, Kyungpook National University, 80 University Road, Buk-gu, Daegu 41566, Korea; jsaeom@knu.ac.kr

³ Department of Mathematics, Natural and Economic Sciences, Ulm University of Applied Sciences, 89075 Ulm, Germany

* Correspondence: stephan.schluter@thu.de

Received: 5 October 2020; Accepted: 16 November 2020; Published: 17 November 2020



Abstract: The Kaya identity is a powerful index displaying the influence of individual carbon dioxide (CO₂) sources on CO₂ emissions. The sources are disaggregated into representative factors such as population, gross domestic product (GDP) per capita, energy intensity of the GDP, and carbon footprint of energy. However, the Kaya identity has limitations as it is merely an accounting equation and does not allow for an examination of the hidden causalities among the factors. Analyzing the causal relationships between the individual Kaya identity factors and their respective subcomponents is necessary to identify the real and relevant drivers of CO₂ emissions. In this study we evaluated these causal relationships by conducting a parallel multiple mediation analysis, whereby we used the fossil fuel CO₂ flux based on the Open-Source Data Inventory of Anthropogenic CO₂ emissions (ODIAC). We found out that the indirect effects from the decomposed variables on the CO₂ flux are significant. However, the Kaya identity factors show neither strong nor even significant mediating effects. This demonstrates that the influence individual Kaya identity factors have on CO₂ directly emitted to the atmosphere is not primarily due to changes in their input factors, namely the decomposed variables.

Keywords: ODIAC; GOSAT; CO₂ flux; mediator analysis; Kaya identity; causality

1. Introduction

Reducing carbon dioxide (CO₂) emissions is currently one of the international community's main goals. However, despite the clarity of this objective, determining which measures have the strongest effects on CO₂ emissions is difficult. To design appropriate strategies, policy makers need detailed insight into causality as well as the effects of individual measures. The Kaya identity is a tool to analyze the drivers of CO₂ emissions by providing a conceptual framework to characterize those driving forces [1]. In this context, CO₂ emissions are disaggregated into five factors: CO₂ divided by fossil fuel-based energy consumption (EC), EC divided by total energy consumption (TEC), TEC divided by the gross domestic product (GDP), GDP divided by population (P), and P itself [2]. Although simple, the Kaya identity has a powerful ability to track progress in implementing CO₂ emission reduction efforts, because many countries already express their climate policies in terms of Kaya components [3–5]: Under the Paris Agreement, treaty parties set nationally determined contribution (NDCs) plans based on the Kaya identity with respect to various CO₂ emission scenarios. Thereby, the parties of the United Nations Framework Convention on Climate Change (UNFCCC) branch out the Kaya identity factors into different sectors to design detailed NDC goals. In this context, policy

makers focus on reducing the driving forces behind Kaya identity factors by managing the intensity of decomposed variables (e.g., TEC and GDP). A common climate change policy is to promote electricity generation based on solar and wind power. To supersede and reduce fossil fuel electricity generation, solar and wind power generation have been growing at about 37% (solar) and 23.4% (wind) per year on average [6].

However, a notable drawback of the Kaya identity is that it is an accounting equation and the driving forces it addresses are not independent. The fixed links of the Kaya identity do not represent causality and ignore non-monotonic or non-proportional effects between individual factors [7]. For instance, scenario builders often assume that high economic growth rates result in high capital turnover, which encourages the development of more advanced and more efficient technologies, in turn leading to lower energy intensities for the economy [2]. Another recent example of dependencies between Kaya identity factors and CO₂ emissions is the decrease in CO₂ emissions induced by the COVID-19 crisis. The COVID-19 pandemic has considerably transformed energy demand patterns around the world by, for example, closure of international borders, underutilization of labor and capital, increased international trade costs, decreased travel services, and a redirection of demand away from activities [8]. As a result, global GDP is predicted to decrease by 6% for single-hit scenarios and by about 7.6% in case of a double-hit scenario, which became reality in the second half of 2020 [9]. In concert with a severe economic recession, by early April 2020 daily global CO₂ emissions declined by 17% relative to 2019 [10].

Most previous studies have focused on the crucial drivers contributing to inventory CO₂ emissions by disaggregating the Kaya identity factors using either the Stochastic Impacts by Regression on Population, Affluence, and Technology (STIRPAT) model or the Logarithmic Mean Divisia Index (LMDI) method [11,12]. LMDI, which applies the Divisia index and logarithmic mean weighted values to decompose CO₂ emissions, quantifies the relative influence of various factors on the changes of an aggregate CO₂ emissions indicator. The method implies that changes in the relative size of factors can change the factors' relative influences on the CO₂ emissions [13,14]. The STIRPAT model, developed as a stochastic version of the impact by population, affluence, and technological development (IPAT) approach, allows testing of the hypothesis that impacts have a weaker relationship with affluence or economic development over time. The STIRPAT model allows the application of sociological theory regarding anthropogenic drivers of environmental impacts, and individual driving forces can be quantified by their respective coefficients [15,16]. Both LMDI and STIRPAT focus on the individual driving forces of the Kaya identity factors, which allows examination of the direct connection between each Kaya identity factor and CO₂ emissions. However, possible interactive effects between the individual Kaya identity factors (and/or their decomposed variables) are neither analyzed nor quantified. Causality issues are also ignored. Han et al. (2018) considered this aspect by conducting a Granger causality test. They intended to explore the directional causality between CO₂ emissions, economic growth, urbanization, and material stocks [17]. Duro and Padilla (2006) proposed applying the Theil index to decompose international inequalities in per capita CO₂ emissions into Kaya identity factors and two interaction terms. They found that the international inequality in per capita CO₂ emissions can be attributed to inequality in per capita income levels [18]. Lately, Hwang et al. (2020) [19] looked closer at the mutual interdependencies between the Kaya identity factors as well as their decomposed variables and the ODIAC-based fossil fuel CO₂ flux. Thereby, ODIAC stands for Open-Source Data Inventory of Anthropogenic CO₂ emissions. Their focus was on EU countries, and they identified different spatial patterns in the computed correlation values. Also, the authors found significant multicollinearity among the decomposed variables, which is not reflected in the Kaya identity itself. Thereby the focus was on the potentials and constraints of the decomposed variables in explaining the variations of intensity of fossil fuel CO₂ flux. However, so far, it has not been deeply explored whether there are any a priori causal relations and indirect effects among the driving forces of Kaya identity factors themselves, or between Kaya identity driving forces such as their subcomponents and actual CO₂ emissions.

Therefore, in this paper, we analyze these interactive effects and causalities in depth using a concept called mediation analysis. Mediation analysis is applied to identify the underlying relations between a dependent variable and an explanatory variable via a third variable, which is called a mediator. This is done by designing and calibrating multivariate linear models in which the mediator variable is inserted between the original independent variables and the dependent variable. It is a fairly simple setup, but it allows understanding of the nature of dependence between dependent variables and explanatory variables via quantifying indirect effects among the explanatory variables and direct effects between individual independent and dependent variables. Hence, it is apt for our purpose: we want to find out tangible evidence for interdependencies among the Kaya identity factors. Can we quantify the causal relations among the Kaya identity factors and their decomposed variables? Can we also quantify the causal effects of the mentioned factors on actual CO₂ emissions? The answers to those questions can help us to monitor the execution effects of mitigation policy efforts and set effective targets for decreasing CO₂ emissions. Our results show—similarly to previous studies [19]—multicollinearity among Kaya identity factors' decomposed variables. Besides which, using mediation analysis, we were able to identify both direct and indirect effects (including causalities) among those variables and from those variables on actual CO₂ emissions, which is a novelty compared to existing studies. We found out that the net (causal) effects among the decomposed variables alleviate the direct effects of their corresponding composite Kaya identity factors on actual CO₂ emissions. This exacerbates to identify the true driving forces among the Kaya identity factors on actual CO₂ emissions. In practical terms: common policy measures like subsidizing electric vehicles aim to influence the decomposed variables of the Kaya identity. However, according to our results, this does not necessary imply a reduction of CO₂ emissions. The Kaya identity factors are only weak and sometimes ambivalent mediators, which implies the need for measures that directly reduce CO₂ emissions.

The paper is structured as follows: in Section 2 the composition of our data set is explained, while, in Section 3, the reader can find an introduction to the methodology used in this paper, with a special focus on mediation analysis. The results of our research are presented in Section 4, and discussed more in detail in Section 5. Here, we comment also on possible implications for politics. Section 6 summarizes and concludes the paper.

2. Materials

There is a wide range of economic and environmental data available; in Section 2.1 we motivate why we focus on a subset of European countries. Section 2.2 contains a details explanation of our set of historic CO₂ observations, while Section 2.3 defines the Kaya identity in detail, gives an overview over all data sources and provides some fundamental statistics for all data sets.

2.1. Countries Included in the Study

European countries in the Annex 1 group have accurate and developed statistical technologies for recording both EC and CO₂ emissions [20,21]. Europe is the second smallest continent in the world after Australia, consisting of 44 highly diverse and densely populated countries. These countries differ with respect to EC, industry, and many other factors [22], making Europe ideal for exploring the causal relations between fossil fuel CO₂ flux and Kaya identity factors under different conditions. In addition, Europe is taking a leading role in implementing active climate change policy, as all countries jointly decided to reduce CO₂ emissions by 40% until 2030 and by even 55% until 2050 (in relation to the 1990 levels). The latter goal was set in the context of the European Green Deal [23], which—among other things—involves the transition of the EU's economies towards using clean, affordable, and secure energy. To have a legal measure for enforcing this goal, i.e. climate neutrality, by 2050, the European Commission designed the European Climate Law [23,24]. Thus, Europe shows more substantial changes with respect to the Kaya identity and CO₂ emissions than other continents. To process the fossil fuel CO₂ flux, accurate preliminary CO₂ emissions data are required because CO₂ emissions are recorded and fixed by in situ surveys. The uncertainty report for Annex 1 countries notes that the

uncertainty of CO₂ emissions associated with fossil fuel combustion is less than 6% on average [20]. European countries are a reliable study area for exploring the causality between decomposed variables and Kaya identity factors and fossil fuel CO₂ flux. For this purpose, we selected 30 Annex 1 countries in Europe, excluding countries that are either too small or too far away from continental Europe. We excluded those countries because it is hard to establish a precise country-specific flux data base as the satellite-based data are recorded on a 1° × 1° grid.

2.2. ODIAC-Based CO₂ Emission Data

Previous literature used inventory net CO₂ emissions on a national scale [25–27]. Net inventory CO₂ emissions fluctuate depending on several variables, such as the collection and reporting system used for the country's energy statistics, data definitions and data processing methods, level of detail, and specific local conditions. In addition, the accuracy, transparency, and uncertainty of inventory CO₂ emission data varies between countries due to differences in the proficiency in and level of development of statistics [28–32]. These fluctuations and uncertainties preclude the interpretation of interactive effects and causality between decomposed variables and Kaya identity factors against CO₂ emissions in reality. To explore the true causality and existing interactive effects between the decomposed variables and Kaya identity factors, as well as their effects on actual CO₂ emissions, we need reliable CO₂ emissions data measured using standardized tools and methods on a regional scale with high spatial resolutions. For this purpose, we chose ODIAC data. ODIAC was developed by Oda and Maksyutov in the context of the Greenhouse Gases Observing Satellite (GOSAT) project at the National Institute for Environmental Studies (NIES) in Japan [33]. The purpose of the ODIAC is to provide precise prior fossil fuel CO₂ emissions data for global and regional CO₂ inversions using the column-averaged CO₂ (XCO₂) data collected by GOSAT. The ODIAC has been widely utilized for the inversion of the official GOSAT Level 4 CO₂ flux data [34–36]. ODIAC data are global, high-resolution monthly emission data, which are the result of a spatial disaggregation of total national emission estimates taken from the Carbon Monitoring for Action (CARMA) data set. It includes emission levels of all types of power plants (fossil fuel, nuclear, hydro, and other renewable energy plants) for over 50,000 locations [37]. ODIAC data at country scale are built to match data from the Carbon Dioxide Information Analysis Center (CDIAC) with a 1 × 1 km spatial resolution [36,38–40]. The ODIAC-based fossil fuel CO₂ flux has detailed information on regional CO₂ sources in terms of distribution patterns and actual fossil fuel combustion with a high spatial resolution, including the international bunker, which is not presented in the Carbon Dioxide Information Analysis Center (CDIAC) [36]. Thus, we can specifically measure the CO₂ amounts directly emitted to atmosphere precisely in those areas where CO₂ sources are located [39,41,42]. Therefore, the ODIAC-based fossil fuel CO₂ flux can facilitate the exploration of realistic causality and interactions among Kaya identity subcomponents against directly emitted to the atmosphere on a regional scale. In this study, we used the ODIAC-based fossil fuel CO₂ flux from GOSAT L4A global CO₂ flux V02.06 provided by the National Institute for Environmental Studies (NIES) of Japan [43].

2.3. Kaya Identity—An Introduction to the Methodology and a Description of Our Data Set

2.3.1. Kaya Identity—The Concept

The IPAT model ($I = P \times A \times T$) explains the level of environmental impact (I) from anthropogenic activities with population size (P), affluence (A)—i.e. the level of income, and technological development (T). Kaya identity is based on the IPAT model, but it is more specific regarding the total level of CO₂ emissions. Kaya identity expresses the CO₂ emissions as the product of four factors: (1) population (P), (2) GDP per capita (GDP/P), (3) energy intensity per unit of GDP (total energy consumptions/GDP), and (4) carbon intensity (CO₂ emissions/total energy consumptions) [44] (Figure 1). Recently, many studies extended the carbon intensity in the traditional Kaya identity by decomposing it into CO₂ emission intensity related to fossil fuel consumption and the share of fossil fuel in the total

energy consumption [11,25,45]. In its extended form, the Kaya identity splits fossil fuel CO₂ emissions into five different factors: P , $G = \text{GDP}/P$, $E = \text{TEC}/\text{GDP}$, $M = \text{EC}/\text{TEC}$, and $I = \text{CO}_2 \text{ emissions}/\text{EC}$. The factors G , E , and M quantify per capita income, energy intensity, and the share of fossil fuel consumption relative to TEC, respectively. The factor I describes the fossil carbon intensity of energy [46]. The corresponding formula reads as follows:

$$\begin{aligned} \text{CO}_2 \text{ emissions} &= \text{Population} \times \frac{\text{GDP}}{\text{Population}} \times \frac{\text{TEC}}{\text{GDP}} \times \frac{\text{EC}}{\text{TEC}} \times \frac{\text{CO}_2 \text{ emissions}}{\text{EC}} \\ &= P \times G \times E \times M \times I \end{aligned} \quad (1)$$

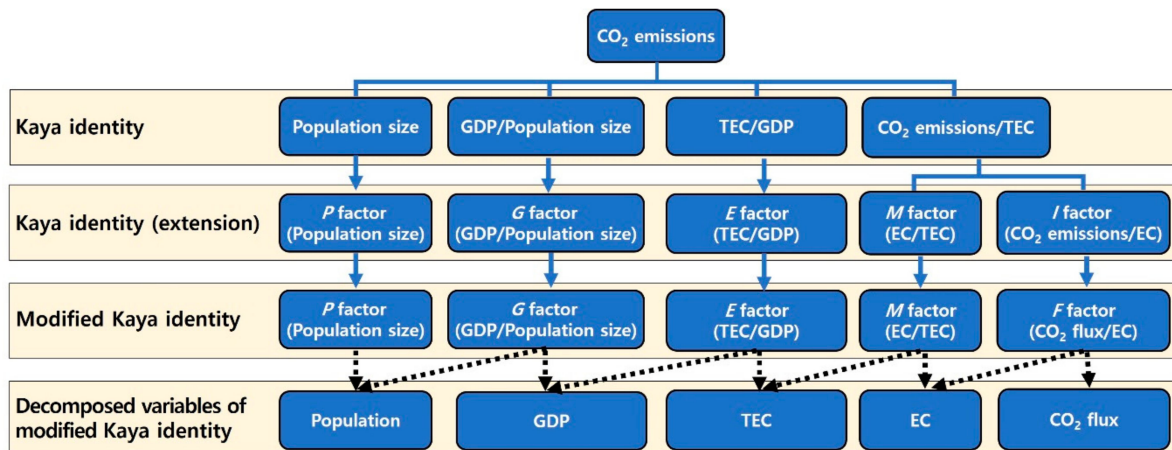


Figure 1. A hierarchy of Kaya identity representations including the decomposed variables.

In a second step we also modify the right side of Equation (1). In order to assess the influences of anthropogenic drivers on fossil fuel CO₂ flux presented in the Kaya identity model, we use a modified form of the Kaya identity with respect to the fossil fuel CO₂ flux [47]. The modified Kaya identity splits up the fossil fuel CO₂ flux multiplicatively instead of the CO₂ emissions. As suggested by Le Quéré et al. (2018) [48], we recalculate the fossil fuel CO₂ flux in CO₂ equivalents (F_{Foss}) to serve as estimates of CO₂ emissions instead of inventory CO₂ emissions in Equation (2). Most factors like P , G , E , and M are analogously defined. For the left side of Equation (1) we follow Raupach et al. (2008) [47] and replace it by $F = F_{\text{Foss}}/\text{EC}$, i.e., the fossil carbon intensity of energy. Hence, in Equation (2), F_{Foss} appears on both sides of the equation.

$$\begin{aligned} F_{\text{Foss}} &= \text{Population} \times \frac{\text{GDP}}{\text{Population}} \times \frac{\text{TEC}}{\text{GDP}} \times \frac{\text{EC}}{\text{TEC}} \times \frac{F_{\text{Foss}}}{\text{EC}} \\ &= P \times G \times E \times M \times F \end{aligned} \quad (2)$$

All model variations and decomposition levels are summed up in Figure 1, which is based on a similar representation in Peters et al. (2017) [49].

2.3.2. Kaya Identity—The Data Sets

The International Energy Agency (IEA) collects EC and production data globally [50,51] and categorizes energy statistics into 13 energy types. In the IEA Energy Statistics [50,51] observations per energy type are converted into energy units (Ktoe). Here, we applied the total final consumption sector from the IEA energy balance from 2010 to 2017 to calculate TEC and EC, which we used to compute the factors E (TEC/GDP), M (EC/TEC), and F factor ($F_{\text{Foss}}/\text{EC}$). To attain EC, we extracted the fossil fuel consumptions (coal, crude oil, oil products, natural gas) from the IEA energy balance. We use 2010–2017 GDP and population data from the World Bank data base in order to derive the factors G ($\text{GDP}/\text{Population}$), E (TEC/GDP), and P (population size). For an overview over variable definitions and the sources of our data sets, see Appendix A.

To compare the factors' variation, we calculated the coefficient of variation (CV), which is defined as a factor's standard deviation divided by its mean. Hence, CV indicates the size of variation relative to a factor's average level. As shown in Table 1, most factors have large CV values. M, which is the share of fossil fuel consumption in net EC shows the lowest CV (approximately 18%), meaning that over the years and over all countries, the ratio of fossil fuel consumption is relatively constant compared to other factors. Apart from that, high CV values are not surprising as we used data from 30 heterogeneous countries.

Table 1. Descriptive statistics for fossil fuel CO₂ flux in CO₂ equivalents (F_{Foss}), five factors of the Kaya identity (E, G, P, F, and M), and the corresponding decomposed variables for 30 European countries from 2010 to 2017.

Category		Min	Max	Mean	STDEV	CV (%)
Kaya identity factors	E (ktoe/MM U.S. \$)	0.03	0.56	0.11	0.09	80.55
	G (1000 U.S. \$/person)	2.02	102.91	31.71	22.25	70.18
	P (MM person)	1.31	82.66	21.67	24.50	113.03
	F (Kt CO ₂ -Equation/ktoe)	2.24	13.28	5.21	2.17	41.64
	M (ktoe)	0.33	0.80	0.6	0.11	18.23
E = TEC/GDP, G = GDP/Population, P = population size, F = F_{Foss} /EC, and M = EC/TEC.						
Decomposed variables of Kaya identity factors	GDP (100 billion U.S. \$)	0.20	38.84	6.53	9.05	138.53
	Population (MM person)	1.31	82.66	21.67	24.50	113.03
	TEC (Mtoe)	2.84	228.90	44.99	51.62	114.73
	EC (Mtoe)	1.29	158.23	30.10	36.62	121.64
	F_{Foss} (10 Mt CO ₂ -Equation yr ⁻¹)	7.14	827.82	138.79	172.65	124.40

Min: Minimum Max: Maximum Mean: Average CV: Coefficients of variation STDEV: Standard deviation.

3. Methodology

Our analysis is based on the calibration results of a multiple linear regression model combined with a mediation analysis. Multiple linear regression is widely known and often used; hence we keep Section 3.1, which concerns the regression models, short and focus on mediation analysis in Section 3.2.

3.1. Multiple Regression Models

To examine the mutual interdependencies between the Kaya identity factors and F_{Foss} , we established a regression model that uses Kaya identity factors as independent variables and F_{Foss} as the dependent variable. Equations (3) and (4) display the corresponding multiple regression models calibrated using the ordinary least squares method.

$$F_{\text{Foss}} = \alpha_0 + \alpha_1 P + \alpha_2 G + \alpha_3 E + \alpha_4 M + \alpha_5 F + \epsilon_1, \quad (3)$$

$$F_{\text{Foss}} = \beta_0 + \beta_1 \text{Population} + \beta_2 \text{GDP} + \beta_3 \text{TEC} + \beta_4 \text{EC} + \epsilon_2, \quad (4)$$

where $\alpha_i, \beta_j \in \mathbb{R}$, $i, j = 1, \dots, 5$, and both ϵ_1 and ϵ_2 are Gaussian distributed with mean values of zero and standard deviations $\sigma > 0$. Equation (3) is the regression model based on the Kaya identity and Equation (4) is the regression model using the decomposed variables of the Kaya identity factors. There are eight years of data available (2010–2017); thus, we have 240 observations for each variable as we aggregate over years and countries. The decomposed variables used in Equations (3) and (4) are annual net amounts and observations of F_{Foss} are annual mean values.

3.2. Mediation Analysis—Introduction and Research Design

Mediation analysis is a statistical method to analyze how the causal antecedent X transmits its effect to the consequent variable Y [52]. A mediation model intends to identify and explain the causality

between an independent variable and a dependent variable by inserting a third variable known as a mediator variable [53]. In lieu of a direct causal relation between the independent and dependent variables, a mediation model suggests that the independent variable affects the mediator variable, which again influences the dependent variable. Thus, the mediator variable illuminates the level of correlation (or interdependence in general) between independent and dependent variables [54]. Mediation analysis helps with finding a valid interpretation of the relationship between independent and dependent variables when those variables seem to have an ambiguous connection.

A parallel multiple mediator model with k mediators has $k + 1$ consequent variables (one for each of the k mediators M_i , $i = 1, \dots, k$, and one for the outcome variable Y). This requires the calibration of $k + 1$ equations to cover all possible effects of X on Y [55], and the formulas for these calibrations are expressed in Equations (5) and (6):

$$M_i = \text{intercept}_{M_i} + a_i X + e_{M_i} \text{ for all } i = 0, \dots, k, \quad (5)$$

$$Y = \text{intercept}_Y + c'X + \sum_{i=0}^n b_i M_i + e_Y. \quad (6)$$

In these parallel multiple mediator equations, a_i measures the effect of X on M_i ; b_i estimates the effect of M_i on Y , controlling X and the other $k - 1$ mediator variables M_i . The factor c' calculates the effect of X on Y , holding all n mediator variables M_i constant [55]. Both intercept_{M_i} and intercept_Y are constant factors. If M_i is an effective mediator, it is substantially correlated with X due to the path from X to M_i (path a). If X explains most of the variation of M_i , there would be no unique variation in M_i to explain Y . According to literature the minimum required sample size for testing coefficients b ($M_i \rightarrow Y$) and c' ($X \rightarrow Y$) is $k(1 - r^2)$, where k is the sample size and r is the correlation coefficient between the causal variable and the mediator. If M_i is a successful mediator (i.e., path or coefficient a_i is large), in order to properly examine coefficients b and c' , we need a comparably larger sample size in order to achieve a test power equivalent to the case of M_i being a weak mediator [56].

In this study, we applied the parallel multiple mediator model to explore the causality between each of the Kaya identity factors as well as their respective decomposed variables and F_{Foss} . Each of the total effects from antecedent X (Kaya identity factors and decomposed variables) on consequent Y (F_{Foss}) could be partitioned into direct and indirect effects through at least one mediator. The proposed mediators (i.e., Kaya identity factors and decomposed variables) can be associated with an outcome either because it gives rise to the outcome or because it correlates with another variable that causally affects the outcome. Hence, including multiple mediators between independent and dependent variables allows detection of the underlying causality. To identify indirect effects among the Kaya identity factors and their decomposed variables, we used all of them as mediators and performed multiple analyses simultaneously. However, in order to do that we had to overcome one problem: our sample size is comparably small as we are handling annual data. Preacher and Hayes (2004) proposed a solution by applying the widely used bootstrapping method [57]. Bootstrapping is a method especially developed for small sample sizes that has the advantage of being non-parametric; therefore, we do not need to impose any distributional assumption on the original data. By resampling the existing set of observations this method allows us to increase the sample size, which, in turn, increases the chance of finding statistically significant mediation effects. For example, if we obtain a non-zero-point estimate, we can only confirm its statistical significance if the zero is not contained in the corresponding confidence interval, the size of which directly depends on sample size because larger samples have smaller confidence intervals. This study has a relatively small number of observations for each variable (240). Hence, by applying the bootstrapping method we increased the chances of finding significant mediation effects. Here we applied the Preacher and Hayes bootstrapping method with 5000 bootstrap samples to explore possible mediation effects.

4. Results

4.1. Two Multiple Linear Regression Models for the Fossil Fuel CO₂ Flux-Based CO₂ Emissions

We calibrated the models from Equations (3) and (4) and performed a mediation analysis (see Section 3.2) to identify causalities between the Kaya identity factors and their decomposed variables.

As Table 2 shows, the multivariate linear regression models based on Equations (3) and (4) have a relatively high explanatory power, with respective R^2 values of 0.85 and 0.95. Both the decomposed variables of Kaya identity factors (Decomposition 1) and the five Kaya identity factors (Decomposition 2) can well explain F_{Foss} . Considering the p -values of the regression coefficients, we see that the coefficients of population and E are statistically insignificant in both decompositions. Hence, their influence on F_{Foss} cannot be proven using the model of Equation (3). Another remarkable result is that Decomposition 2 shows significant multicollinearity whereas Decomposition 1 does not. Similar results were demonstrated by Hwang et al. (2020) [19], which—contrary to this analysis—only focused on the superficial relations between the decomposed variables and the ODIAC-based fossil fuel CO₂ flux with an additional focus on the country-specific situation. Here we look closer into the causal relations between the individual factors by using mediation analysis (see Section 4.2).

Table 2. Results of the multivariate linear regression of (a) decomposed variables on F_{Foss} and (b) Kaya identity factors on F_{Foss} .

Category		Standardized Coefficient	t-Statistics	p-Value	Variance Inflation Factor (VIF)
Decomposed variables of five factors in Kaya identity	Population	−0.11	−2.31	0.22	9.30
	GDP	−0.43	−7.98	0.00	12.79
	TEC	0.73	4.69	0.00	108.62
	EC	0.74	4.58	0.00	116.37
	(a) R: 0.973 R ² : 0.947 Durbin-Watson: 2.11 F-value (p-value): 1055.64 (0.00)				
Kaya identity factors	E	0.02	0.54	0.59	1.63
	G	0.18	4.97	0.00	2.05
	P	0.93	31.76	0.00	1.32
	F	0.19	6.05	0.00	1.50
	M	0.06	1.97	0.05	1.34
(b) R: 0.921 R ² : 0.849 Durbin-Watson: 1.849 F-value (p-value): 263.31 (0.00)					

E = TEC/GDP, G = GDP/Population, P = population size, F = $F_{\text{Foss}}/\text{EC}$, and M = EC/TEC.

Note that the factor Population, which is identical to Kaya identity factor P, is not statistically significant (with a p -value of 0.22) whereas P is (p -value of 0). Hence, all relevant information about population size must already be included in the other factors. Thus, we also tested various multiple regression models with different combinations of decomposed variables from Decomposition 1. The results of this test are shown in Table 3. We saw that Population is statistically significant when it is used as an exogenous variable in combination with GDP, TEC, or EC. However, if two of the variables are included in the regression, the p -value of Population increases significantly (up to 0.34). Regarding multicollinearity between GDP, TEC, and EC, the VIF values vary with the model setup. If at least two of GDP, TEC, or EC are used together in the regression, we saw a considerable increase of variance inflation factor (VIF) values (up to 26.23). These VIF values are remarkably high when TEC and EC are used conjointly in the regression model. This seems reasonable when considering the fairly high correlation value between both variables (see Table 4). To look more deeply into the interdependencies among Decomposition 1 and 2, we applied mediation analysis as described in Section 3.2, and use F_{Foss} as endogenous variable in all setups.

Table 3. Results of different multiple regression models with different combinations of independent variables (decomposed variables of the Kaya identity) and F_{Foss} as a dependent variable.

Category	Standardized Coefficient	t-Statistics	p-Value	VIF
Population	0.56	13.34	0.00	3.28
GDP	0.41	9.88	0.00	3.28
R: 0.93 R ² : 0.87 Durbin-Watson: 2.13 F-value (p-value): 811.81 (0.00)				
Population	0.08	1.75	0.08	7.14
TEC	0.89	19.58	0.00	7.14
R: 0.97 R ² : 0.93 Durbin-Watson: 1.99 F-value (p-value): 1606.50 (0.00)				
Population	0.06	1.24	0.22	7.50
EC	0.91	19.45	0.00	7.50
R: 0.97 R ² : 0.93 Durbin-Watson: 2.12 F-value (p-value): 1591.92 (0.00)				
Population	−0.08	−1.74	0.08	9.19
GDP	−0.38	−6.90	0.00	12.30
EC	1.40	16.90	0.00	28.11
R: 0.97 R ² : 0.94 Durbin-Watson: 2.06 F-value (p-value): 1285.93 (0.00)				
Population	−0.04	−0.95	0.34	8.48
GDP	−0.37	−6.80	0.00	12.05
TEC	1.35	16.95	0.00	26.23
R: 0.97 R ² : 0.94 Durbin-Watson: 2.06 F-value (p-value): 1291.13 (0.00)				
Population	0.06	1.19	0.24	7.50
TEC	0.49	2.82	0.01	104.41
EC	0.43	2.43	0.02	109.60
R: 0.97 R ² : 0.93 Durbin-Watson: 2.05 F-value (p-value): 1095.20 (0.00)				

Table 4. Correlation between Kaya identity factors and their decomposed variables.

Category	Population	Decomposed Variables			Kaya Identity Factors					
		GDP	TEC	EC	F_{Foss}	E	G	P	F	M
Population	1	0.83 **	0.93 **	0.931 **	0.91 **	−0.01	−0.063	1.00 **	−0.18 **	0.47 **
GDP	−	1	0.95 **	0.95 **	0.88 **	−0.34 **	0.286 **	0.83 **	−0.29 **	0.38 **
TEC	−	−	1	0.99 **	0.97 **	−0.14 *	0.115	0.93 **	−0.23 **	0.42 **
EC	−	−	−	1	0.97 **	−0.15 *	0.097	0.93 **	−0.22 **	0.48 **
F_{Foss}	−	−	−	−	1	−0.04	0.008	0.91 **	−0.08	0.44 **
E	−	−	−	−	−	1	−0.61 **	0.01	0.37 **	−0.08
G	−	−	−	−	−	−	1	−0.06	−0.52 **	−0.09
P	−	−	−	−	−	−	−	1	−0.18 **	0.47 **
F	−	−	−	−	−	−	−	−	1	−0.16 *
M	−	−	−	−	−	−	−	−	−	1

*: $p \leq 0.05$; **: $p \leq 0.01$.

4.2. Interactions between the Decomposed Variables and Kaya Identity Factors and the ODIAC-Based Fossil Fuel CO₂ Flux Detected through Mediation Analysis

4.2.1. Indirect Effects Among the Decomposed Variables on the Fossil Fuel CO₂ Flux-Based CO₂ Emissions

Table 5 shows mediation analysis results for all Decomposition 1 variables. Confidence intervals are computed based on a 5% error probability. We see that Population strongly influences F_{Foss} through GDP, TEC, and EC. The indirect effect of Population through GDP, TEC, and EC on F_{Foss} varies from −39.36% to 76.10%, meaning that the other decomposed variables strongly dominate the relationship between Population and F_{Foss} . There is seemingly no direct relationship between Population and F_{Foss} ; however, there is an indirect one that is mediated by some decomposed variables (Figure 2). The strong mediation effect causes Population to be statistically insignificant in the regression on F_{Foss} . In contrast, Population has a comparatively weaker mediation effect on GDP, TEC, and EC, ranging from −10.19% to −9.99%. However, GDP, TEC, and EC have strong mediating effects on each other, with values ranging from −42.09% to 79.69 % on F_{Foss} (see Table 5). This means that GDP, TEC, and EC are highly

dependent on each other, which supports the results of the above correlation analysis (see Table 4). When GDP, TEC, and EC are applied as mediators to each other, they mediate the other decomposed variables by explaining most of the variation caused by mediated variables. Thereby, the high indirect effects from GDP, TEC, and EC on each other indicate high multicollinearity among these variables.

Table 5. Mediation analysis results among the decomposed variables of all Kaya identity factors on F_{Foss} .

Category	Mediator	R	R ²	p-Value	Coeff.	Std. Coeff.	Indirect Effect		
							Indirect (%)	LLCI	ULCI
Population	GDP	0.83	0.70	0.00	0.31	0.83	−39.36	−3.50	−1.69
	TEC	0.93	0.86	0.00	1.95	0.93	74.92	2.33	7.76
	EC	0.94	0.87	0.00	1.39	0.93	76.10	1.90	8.10
	F_{Foss}	0.91	0.82	0.00	6.38	0.91	-	-	-
Total effect (p-value): 6.38 (0.00) Direct effect (p-value): −0.74 (0.12)									
GDP	Population	0.83	0.70	0.00	2.26	0.83	−9.99	−3.96	0.64
	TEC	0.95	0.90	0.00	5.42	0.95	78.81	6.22	21.19
	EC	0.95	0.90	0.00	3.84	0.95	79.69	5.20	22.04
	F_{Foss}	0.88	0.78	0.78	16.83	0.58	-	-	-
Total effect (p-value): 16.82 (0.00) Direct effect (p-value): −8.16 (0.00)									
TEC	Population	0.93	0.86	0.00	0.44	0.93	−10.15	−0.76	0.13
	GDP	0.95	0.90	0.00	0.17	0.95	−42.09	−1.84	−0.95
	EC	1.00	0.99	0.00	0.71	1.00	76.38	0.98	4.06
	F_{Foss}	0.97	0.93	0.93	3.23	0.97	-	-	-
Total effect (p-value): 3.23 (0.00) Direct effect (p-value): 2.45 (0.00)									
EC	Population	0.93	0.87	0.00	0.62	0.93	−10.19	−0.76	0.13
	GDP	0.95	0.90	0.00	0.23	0.95	−42.06	−1.84	−0.95
	TEC	1.00	0.99	0.00	1.40	1.00	75.50	0.98	4.06
	F_{Foss}	0.97	0.93	0.00	4.55	0.97	-	-	-
Total effect (p-value): 4.55 (0.00) Direct effect (p-value): 3.49 (0.00)									
F_{Foss}	Population	0.91	0.82	0.00	0.13	0.91	-	-	-
	GDP	0.88	0.78	0.00	0.05	0.88	-	-	-
	TEC	0.97	0.93	0.00	0.29	0.97	-	-	-
	EC	0.97	0.93	0.00	0.21	0.97	-	-	-
Total effect (p-value): 1.00 (0.00) Direct effect (p-value): 1.00 (0.00)									

Coeff.: coefficients; Std. coeff.: standardized coefficients; LLCI: lower limit of the bootstrap confidence interval; ULCI: upper limit of the bootstrap confidence interval; F_{Foss} : fossil fuel CO₂ flux-based CO₂ emissions.

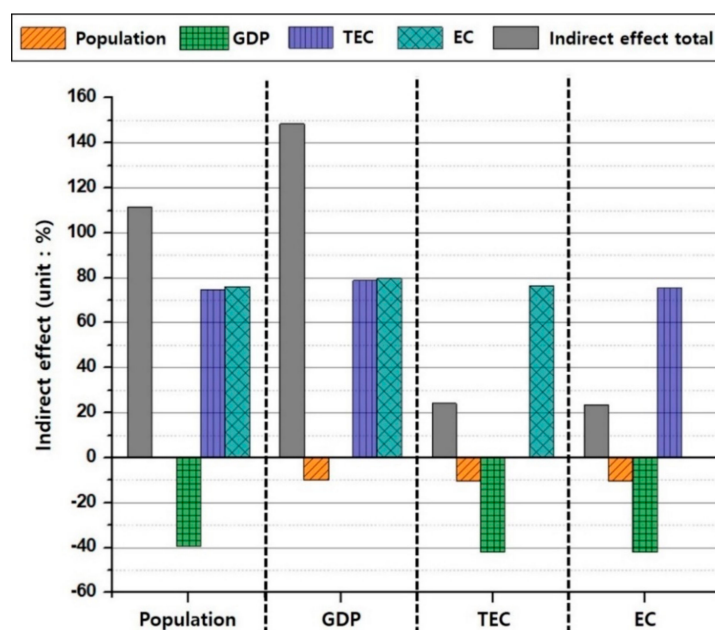


Figure 2. Comparisons of indirect effects among the decomposed variables of Kaya identity factors on F_{Foss} .

In this regard, we also see that the lack of statistical significance of E, which is calculated by dividing TEC by GDP, is also due to the opposite mediating effects between TEC and GDP. A strong correlation between X (TEC, GDP) and the mediator (E factor) can inflate the standard error of Path b (E factor \rightarrow F_{Foss}). If Path b is not significant because of multicollinearity, the indirect effect necessary to create mediation would be probably insignificant [58]. GDP has a significant negative mediating effect on TEC, at a level of -42.09% . This means that GDP depresses the direct effect of TEC on F_{Foss} at a level of -42.09% . In contrast, TEC has a positive mediating effect on GDP, encouraging a direct effect of GDP on F_{Foss} at a level of 78.81% . Hence, TEC and GDP mutually compensate their influences on F_{Foss}, which flattens the influence E has on F_{Foss} and causes the insignificance.

4.2.2. Indirect Effects between the Decomposed Variables and Kaya Identity Factors on ODIAC-Based Fossil Fuel CO₂ Flux

In our mediation analysis we observed that Kaya identity factors have weakly negative mediating effects on the casualty between the decomposed variables of Kaya identity factors and F_{Foss}. Indirect effects of Kaya identity factors range from -5.86% to -2.12% , except for GDP versus Kaya identity factors, which is 31.69 (Table 6). Even though the mediator variables, which are the Kaya identity factors in this subsection, do not change the direction of the relationships between the decomposed variables and F_{Foss}, they slightly decrease the strength of the existing relationships (Figure 3). In particular, the high (total) indirect effect of GDP results from the high indirect effects of P. That is because countries with a comparably higher GDP, such as France, Germany, Italy, Spain, and the United Kingdom, are among the most densely populated countries in this study. In addition, the Kaya identity factors and their respective decomposed variables are statistically insignificant in terms of mediating effects. Table 6 displays statistically insignificant indirect effects between the Kaya identity factors and their decomposed variables in three scenarios: (1) Population \rightarrow G factor \rightarrow F_{Foss}, (2) GDP \rightarrow G factor \rightarrow F_{Foss}, and (3) EC \rightarrow M factor \rightarrow F_{Foss}. Both G and M factor do not play a statistically significant mediating roles for their decomposed variables, which is also true for Population, GDP, and EC. This can be shown by considering the confidence interval for the case where G operates as mediator for Population, which ranges from a lower level of -0.21 to an upper level of 0.05 . The confidence intervals for the scenarios where G operates as a mediator for GDP or M factor as a mediator for EC are $[-0.02, 0.34]$ and $[-0.10, 0.2]$, respectively. In all three cases, zero is included in the confidence interval. Hence the effects are not statistically significant. In other words, the variation of some Kaya identity factors (G and M factor) cannot be explained by their decomposed variables.

Given our dataset, there is a fair chance that we observe significant multicollinearity in the regression of the decomposed variables on F_{Foss} as indicated by the VIF values in Table 2. Mediation analysis reflects these interdependencies as net effects among the decomposed variables are quantitative visible and measurable. As displayed in Table 5 and Figure 2 we see indirect effects between -42.09% and $+79.96\%$. These are significant effects, which, besides multicollinearity, are underestimated or in the Kaya identity model, where a mediation analysis shows weaker indirect effects between -5.26% and 38.29% (see Table 6 and Figure 3). Hence, changes of the decomposed variables have a considerably smaller effect on the CO₂ emissions because of weak mediating effects of the Kaya identity factors and strong mediating effects among the decomposed variables. Quantifying such effects is a benefit of mediation analysis and contributes to the existing literature. Previous studies have already attempted to account for multicollinearity. Authors like Purcel (2020) [59], Georgiev and Mihaylov (2015) [60], and Choi et al. (2010) [61] mainly focused on empirically explaining interdependencies using the environmental Kuznets curve (EKC) [62]. Multicollinearity is sometimes also accounted for by decomposing the independent variables and excluding selected ones from the regression [63]. However, despite being reasonable, these decomposition approaches face the danger of an omitted variable bias. Besides, if you first decompose and then eliminate input factors, the whole theoretic foundation of the regression model itself might be altered [64]. Consider, for example the Kaya identity or the IPAT modification. These are basically composite indices that have been in use for the past

15–20 years, and many studies and political decisions have been based on findings derived from this model. If you decompose the influence factors and eliminate those causing multicollinearity, you have to reconsider all theories based on the Kaya identity model.

Table 6. Results of mediation analysis on Kaya identity factors in a relationship between decomposed variables and F_{Foss} .

Category	Mediator	R	R ²	p-Value	Coeff.	Std. Coeff.	Indirect Effect		
							Indirect (%)	LLCI	ULCI
Population	E	0.01	0.00	0.85	0.00	−0.01	−0.02	−0.06	0.01
	G	0.06	0.00	0.23	−0.06	−0.06	−1.25	−0.21	0.05
	F	0.18	0.03	0.00	−0.02	−0.18	−3.83	−0.35	−0.14
	M	0.45	0.22	0.00	0.00	0.47	2.99	0.08	0.31
	Total effect (p-value): 6.38 (0.00) Direct effect (p-value): 6.52 (0.00)								
GDP	E	0.34	0.11	0.00	0.00	−0.34	−5.26	−1.17	−0.61
	G	0.29	0.08	0.00	0.70	0.29	0.87	−0.02	0.34
	F	0.29	0.09	0.00	−0.07	−0.29	−4.86	−1.11	−0.59
	M	0.38	0.14	0.00	0.00	0.38	2.65	0.26	0.67
	P	0.83	0.70	0.00	2.26	0.83	38.29	5.08	8.10
	Total effect (p-value): 16.82 (0.00) Direct effect (p-value): 11.49 (0.00)								
TEC	E	0.14	0.02	0.00	0.00	−0.14	−0.83	−0.04	−0.01
	G	0.12	0.01	0.00	0.05	0.12	0.04	−0.01	0.01
	F	0.23	0.05	0.00	−0.01	−0.23	−3.31	−0.15	−0.07
	M	0.42	0.17	0.00	0.00	0.42	2.80	0.06	0.13
	P	0.93	0.86	0.00	0.44	0.93	−3.07	−0.43	0.24
	Total effect (p-value): 3.23 (0.00) Direct effect (p-value): 3.37 (0.00)								
EC	E	0.15	0.02	0.00	0.00	−0.15	−1.12	−0.08	−0.03
	G	0.10	0.01	0.01	0.06	0.10	0.18	−0.08	−0.03
	F	0.22	0.05	0.00	−0.01	−0.22	−2.94	−0.19	−0.08
	M	0.48	0.23	0.00	0.00	0.48	−0.87	−0.10	0.02
	P	0.93	0.87	0.00	0.62	0.93	−1.11	−0.52	0.46
	Total effect (p-value): 4.55 (0.00) Direct effect (p-value): 4.81 (0.00)								

Coeff.: coefficients; Std. coeff.: standardized coefficients; LLCI: lower limit of the bootstrap confidence interval; ULCI: upper limit of the bootstrap confidence interval; P = population size; G = GDP/Population; E = TEC/GDP; M = EC/TEC; and F = $F_{\text{Foss}}/\text{EC}$.

Alternatively, some case studies examine the correlation, i.e., interdependencies between CO₂ emission and its contributors with composite index factors that are the result of combining two highly correlated variables. In other words, the composite index is used as a measure to eliminate multicollinearity. Tavakoli [4], for example, conducted a multiple regression on the Kaya identity factors from the top ten CO₂ emitting countries. It postulates that population, energy intensity, and GDP per capita are major influential factors. However, if—as in our data set—the measured correlation is significantly large but not perfect (i.e., ± 1), a model based on composite factors might over- or underestimate the net effects of changes in the original independent variables [65]. Even if this is not the case: a composite model provides no precise insight, whereby our mediation analysis does.

To sum up our results: because it is important for policy makers, there is substantial literature exploring the driving forces of the decomposed variables of Kaya identity. As explained, current approaches have limitations as problems like omitted variable bias or over-/underestimation of a parameter's influence might occur. In contrast, using mediation analysis in this study, we are able to evaluate the interactive net effects among the Kaya identity factors and their decomposed variables on direct CO₂ emissions to the atmosphere. This is especially of relevance as we analyze recent emissions from 30 European countries, which, being classified as post-industrial economies, have entered a new stage in the EKC [66]. These insights can contribute to decisions made by EU policymakers, who currently struggle to find appropriate measures to mitigate CO₂ emissions. Knowledge about the direct and indirect effects of the decomposed factors of Kaya identity provides valuable information for this purpose.

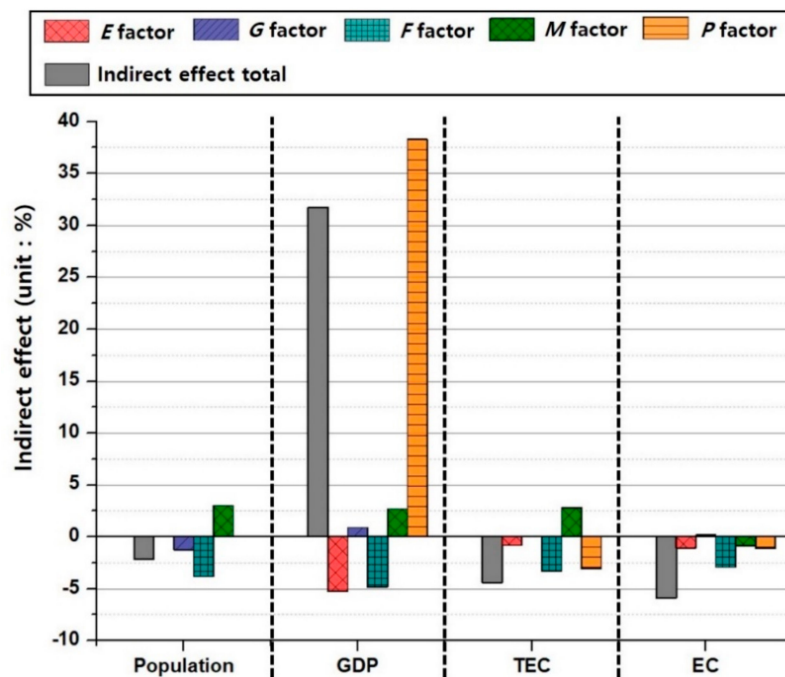


Figure 3. Comparisons of indirect effects of Kaya identity factors in a relationship between decomposed variables and F_{Foss} .

5. Discussion

In Section 5.1, we discuss implications of the multicollinearity and the mediation effects found in Section 4 in more detail. Besides, in Section 5.2, we briefly reflect on the benefits of the Kaya identity concept in the light of both our findings and of the criticism found in contemporary literature. This is important, as we can only draw accurate conclusions from our findings if we are also aware of the model's shortcomings.

5.1. Consequences of Multicollinearity

As noted in Section 4.1, the decomposed variables of the Kaya identity factors show strong multicollinearity. That means changes in the levels of the decomposed variables significantly account for variation in other decomposed variables. Hence, we concluded that the decomposed variables are controlled by other decomposed variables that function as mediating factors in our study. On the contrary, the variance of the Kaya identity factors cannot be substantially explained by their decomposed variables as they are independent of each other. The Kaya identity is by definition a multiplicative identity, and its factors are computed by dividing two decomposed variables. It is thus the ratio and proportional intensity of the anthropogenic CO₂ emitting activities. The calculation process eliminates the multicollinearity of the respective decomposed variables. The results of the mediation model in Table 6 support the Kaya identity factors' independence.

Multicollinearity of the decomposed variables means that if two decomposed variables of a specific factor are changed with similar growth rates, the driving force of the factor on CO₂ emissions will be stable or change only slightly. As TEC, EC, and GDP remain stable or decline, it stands to reason that CO₂ emissions are likely to remain stable or decline as well. To examine this slightly different scenario with a focus on the dynamics, i.e., the relationship between changes in Kaya identity factors and changes in their decomposed variables, we fitted another linear model. This is based on a concept called proportional growth rate $r(X)$, which is defined as $r(X) = \frac{X'}{X}$ (%/year) [47]. We fitted the model to all combinations of growth rates of both decomposed variables and Kaya identity factors, and found that—except for GDP—all models have low R^2 values, ranging between 0.02 and

0.54 (Figure 4). The GDP's R^2 is comparatively high, especially considering the regression of GDP's growth rate on the growth rates of both E and G, with values of 0.84 and 0.99, respectively (Figure 4). This means that—except for GDP, Population, and G—there is only a fairly weak connection between the decomposed variables' proportional growth rates and the corresponding Kaya identity factors' growth rate.

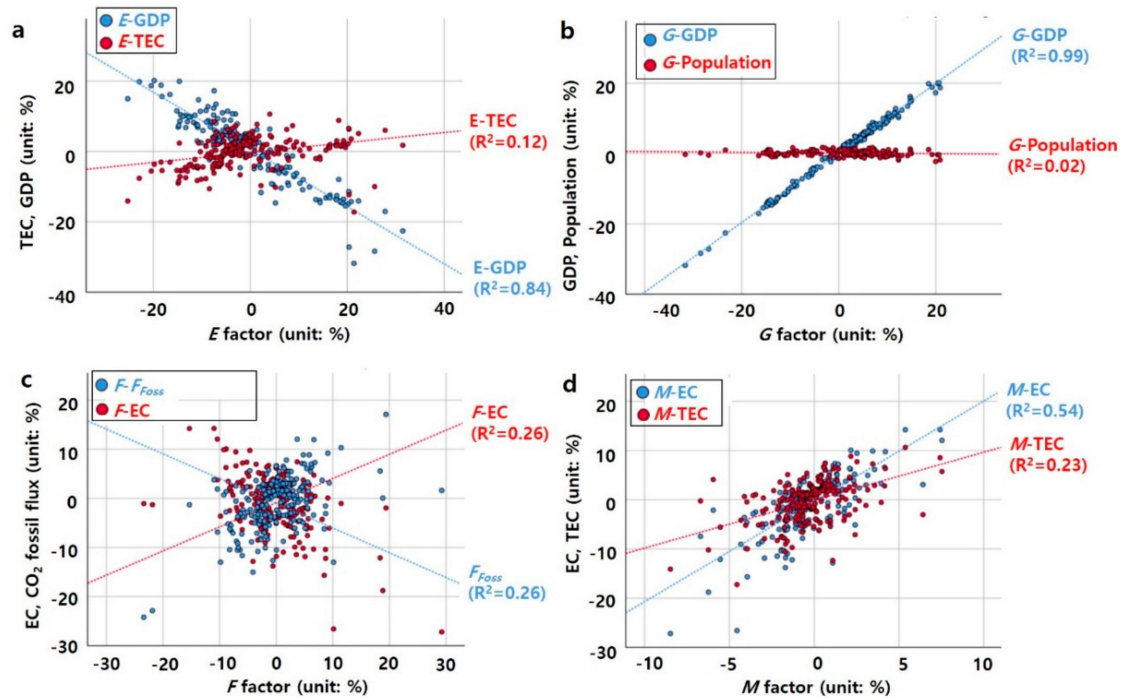


Figure 4. Scatter plots of proportional growth rates of decomposed variables (x axis) and Kaya identity factors (y axis). (a) TEC and GDP versus E, (b) GDP and population versus G, (c) F_{Foss} and GDP versus F, and (d) TEC and EC versus M.

As mentioned, we measured multicollinearity among the decomposed variables as well as strong mediating effects among them. Hence, the decomposed variables have a comparatively weak influence on the Kaya identity factors. In addition, unexpected events such as economic crises, warmer winters, or the COVID-19 pandemic affect the individual decomposed variables simultaneously. For example, net EC and fossil fuel combustion are the major drivers of and targets for mitigating CO_2 emissions. Therefore, reducing the decomposed variables does not always yield the desired effect, as the respective resulting Kaya identity factors may not have a large enough negative effect on CO_2 emissions.

5.2. Kaya Identity in the Context of Contemporary Literature

Fischer-Kowalski and Amann (2001) [67] stated that—in the context of the Kaya identity—population (P) and technology (E, M, F) seem to dominate GDP per capita or affluence (G) in terms of environmental impact (CO_2 emissions). Looking closer, affluence is related to the development and the internal structure of the economic sector, where (besides mobility) the global supply chains of companies are one of the major drivers of CO_2 emissions. Only considering the factor G in the Kaya identity ignores the complex interdependencies among and between the country-specific economic sectors. Ahi et al. (2013) [68], Gong et al. (2019) [69], and Zimon et al. (2019) [70], for example, addressed this problem by analyzing and proposing a green supply chain management. Karl and Ranné (1997) and Sadik-Zada and Gatto (2020) stated that conjoint efforts towards a decarbonization-friendly transition of the energy mix, tertiarization, general trade openness, and green supply chains would significantly contribute to mitigating CO_2 emissions [24,71–73].

Our analysis cannot cover these international developments, as it is based on the Kaya identity. However, by considering the decomposed factors and the interdependencies among them, we provide a basis for further research concerning intercountry effects.

Peters et al. (2017) discussed the recent slowdown of CO₂ emissions with Kaya-derived indicators by using an interconnected and nested structure composed of different forms of Kaya identity components [49]. They observed that economic factors and energy efficiency have contributed more to decreased CO₂ emissions than the adoption of wind and solar power. They showed that most of the indicators are currently consistent with emission scenarios from the Paris Agreement goal of keeping the temperature increase below 2 °C. Nonetheless, literature shows that the lack of carbon capture and storage (CCS) technologies, slow improvements in energy efficiency, or missing transformations of energy structures would threaten the 2030 goals and net-zero emissions targets of the Paris Agreement. These issues are also well-projected in this study. Changes in the decomposed variables cannot directly and effectively affect the reduction of fossil fuel CO₂ flux because of multicollinearity among them. To increase practical CO₂-mitigating effects from key Kaya identity factors, CO₂ emissions must be reduced at the same time. In this regard, additional policy-wise efforts, e.g., towards increasing energy efficiency, reducing energy consumption itself, and/or implementing new technologies like CCS are required. Independent of politics, industry has to contribute as well. e.g., by designing and implementing sustainably strategies (see, e.g., Scarpato et al. (2020) [74]) or, as mentioned above, focusing on a greener, i.e. less carbon-intensive, supply chain.

6. Conclusions

In this study, we evaluated the causal relationship between the Kaya identity factors and their decomposed variables to explore the real interactive influences on F_{FOSS} , which is the CO₂ emitted from in situ fossil fuel CO₂ sources to the atmosphere. We also include indirect effects among the decomposed variables. The analysis was done by evaluating a linear model where the decomposed variables of the Kaya identity serve as explanatory variables and F_{FOSS} as dependent variable. An analogous model was set up for the Kaya identity factors as well. Also, to look deeper into the direct and indirect dependencies between the individual variables, we conducted an extensive mediation analysis.

Results (Section 4) show that these indirect effects are especially significant; however, there are no strong or significant mediating effects from the Kaya identity factors on their decomposed variables. With respect to the proportional growth rate, changes in the decomposed variables do not lead to diverse and enormous growth rate changes of Kaya identity factors. Because of the strong indirect effects among the decomposed variables on F_{FOSS} , policy makers must consider the causalities between the decomposed variables and Kaya identity factors to set realistic carbon mitigation targets. The quantitative net effects (direct effects and indirect effects) among the Kaya identity factors and their decomposed variables indicate that there are hidden and intricate interdependencies.

For example, weak indirect effects between the Kaya identity factors themselves suggest that changes in the decomposed variables unproportionally affect the Kaya identity factors and hence F_{FOSS} . Our analysis using proportional growth rates confirms these findings: except for G , GDP , and population, there are only weak interdependencies between the growth rates of the Kaya identity factors and its corresponding decomposed variables. Furthermore, as discussed in Section 5.2, the Kaya identity model is a national model—the international flow of goods and energy, for example, is not considered.

To sum up our findings: technically speaking, most policy measures target one or more decomposed variables of the Kaya identity [3–5]—if possible while keeping negative effects on the economy at a minimum. Policymakers thereby especially focus on E (CO₂ emissions/EC) and M (EC/TEC). The factor P cannot be controlled that easily, and both G and I are closely related to economic growth, i.e. should not be influenced negatively. Popular examples are subsidies for electric vehicles and renewable energies like solar power. However, given the weak mediating effects of the Kaya identity factors, which are, according to the Kaya model equation, eventually responsible for reducing CO₂ emissions, this is not enough. Even if policy measures to mitigate CO₂ emissions successfully influence either E

or M or both, this does not necessarily imply a reduction of emissions on a national level. The results of our study can be used to account for such unforeseen effects and adjust policy strategies adequately. Besides this, our results show the need for a global sustainable strategy to reduce CO₂ emissions directly and not only by influencing one of the factors discussed above. However, some promising new technologies like CCS are controversially discussed, as their long-term impact on environment is still unclear. A lot of measures that reduce emissions are also quite costly. This shows the need for more research about the cross-border effects of CO₂ emissions reduction measure, e.g., along the global supply chain, to decide which of them is most effective. Besides, note that when interpreting this study, we had a short dataset consisting of observations from eight consecutive years. Additionally, only European countries were included in this study. Hence, more research involving a longer set of observations (20 years or more if possible) and data from other continents is required and would produce more information for designing appropriate policy measures.

Author Contributions: Conceptualization, Y.H. and J.-S.U.; methodology, S.S.; software, Y.H. and J.H.; validation, S.S., J.-S.U., Y.H., and J.H.; formal analysis, J.-S.U.; resources, Y.H. and J.H.H.; data curation, Y.H., S.S. and J.H.; writing—original draft preparation, S.S. and Y.H.; writing—review and editing, Y.H., S.S., and J.-S.U.; visualization, Y.H.; supervision, S.S.; project administration, J.-S.U. and S.S.; funding acquisition, J.-S.U., S.S., and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This article is co-funded by the Baden-Württemberg Ministry of Science, Research and Culture.

Acknowledgments: We thank Japan Aerospace Exploration Agency (JAXA), The Ministry of the Environment, Japan (MOEJ), National Institute for Environmental Studies (NIES), International Energy Agency (IEA), World bank and United Nations Framework Convention on Climate Change (UNFCCC) for providing GOSAT satellite data, energy consumption, population, Gross Domestic Product (GDP) and National Inventory Report data.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

CARMA	Carbon Monitoring for Action
CDIAC	Carbon Dioxide Information Analysis Center
CV	Coefficient of variation
EC	Fossil-fuel based energy consumption
E	CO ₂ emissions/EC
EKC	Environmental Kuznets curve
F	CO ₂ flux/EC
F _{Foss}	ODIAC-based fossil fuel CO ₂ flux
G	GDP/P
GDP	Gross domestic product
GOSAT	Japan Aerospace Exploration Agency Greenhouse Gases Observing Satellite
I	TEC/GDP
IEA	International Energy Agency
M	EC/TEC
NDCs	Nationally determined contributions
LMDI	Logarithmic Mean Divisia Index
ODIAC	Open-Source Data Inventory of Anthropogenic CO ₂ emissions
P	Population size
STIRPAT	Stochastic Impacts by Regression on Population, Affluence and Technology
TEC	Total energy consumptions
UNFCCC	United Nations Framework Convention on Climate Change

Appendix A

Table A1. Descriptions of datasets applied in calculating Kaya identity factors.

Kaya Identity Factor	Decomposed Variables	Dataset
P factor	Population	Population data from the World Bank data
G factor (GDP/Population)	GDP Population	GDP (constant 2010 US\$) from World Bank Population data from the World Bank data
E factor (TEC/GDP)	TEC GDP	Total final consumption from the IEA energy balance GDP (constant 2010 US\$) from the World Bank
M factor (EC/TEC)	EC TEC	Fossil fuel consumptions (coal, crude oil, oil products, natural gas) from the IEA energy balance Total final consumption from the IEA energy balance
F factor (CO ₂ flux/EC)	CO ₂ flux EC	ODIAC-based fossil fuel CO ₂ flux Fossil fuel consumptions (coal, crude oil, oil products, natural gas) from the IEA energy balance

References

- Wang, J.; Hu, M.; Tukker, A.; Rodrigues, J.F.D. The impact of regional convergence in energy-intensive industries on China's CO₂ emissions and emission goals. *Energy Econ.* **2019**, *80*, 512–523. [\[CrossRef\]](#)
- IPCC. *Special Report on Emissions Scenarios*; IPCC: Cambridge, UK, 2000.
- Aye, G.C.; Edoja, P.E. Effect of economic growth on CO₂ emission in developing countries: Evidence from a dynamic panel threshold model. *Cogent Econ. Financ.* **2017**, *5*, 1379239. [\[CrossRef\]](#)
- Tavakoli, A. A journey among top ten emitter country, decomposition of “Kaya Identity”. *Sustain. Cities Soc.* **2018**, *38*, 254–264. [\[CrossRef\]](#)
- UNFCCC. *Synthesis Report on the Aggregate Effect of the Intended Nationally Determined Contributions*; UNFCCC: Bonn, Germany, 2015.
- IEA. *Renewables Information 2019*; IEA: Paris, France, 2019.
- Aguir Bargaoui, S.; Liouane, N.; Nouri, F.Z. Environmental Impact Determinants: An Empirical Analysis based on the STIRPAT Model. *Procedia Soc. Behav. Sci.* **2014**, *109*, 449–458. [\[CrossRef\]](#)
- Maliszewska, M.; Mattoo, A.; van der Mensbrugghe, D. *The Potential Impact of COVID-19 on GDP and Trade: A Preliminary Assessment*; World Bank: Washington, DC, USA, 2020.
- OECD. *Real GDP Forecast (Indicator)*; OECD Data; OECD: Paris, France, 2020. [\[CrossRef\]](#)
- Le Quéré, C.; Jackson, R.B.; Jones, M.W.; Smith, A.J.P.; Abernethy, S.; Andrew, R.M.; De-Gol, A.J.; Willis, D.R.; Shan, Y.; Canadell, J.G.; et al. Temporary reduction in daily global CO₂ emissions during the COVID-19 forced confinement. *Nat. Clim. Chang.* **2020**, *10*, 647–653. [\[CrossRef\]](#)
- Jung, S.; An, K.-J.; Dodbib, G.; Fujita, T. Regional energy-related carbon emission characteristics and potential mitigation in eco-industrial parks in South Korea: Logarithmic mean Divisia index analysis based on the Kaya identity. *Energy* **2012**, *46*, 231–241. [\[CrossRef\]](#)
- Chontanawat, J. Driving Forces of Energy-Related CO₂ Emissions Based on Expanded IPAT Decomposition Analysis: Evidence from ASEAN and Four Selected Countries. *Energies* **2019**, *12*, 764. [\[CrossRef\]](#)
- Kim, S. LMDI Decomposition Analysis of Energy Consumption in the Korean Manufacturing Sector. *Sustainability* **2017**, *9*, 202. [\[CrossRef\]](#)
- Meng, Z.; Wang, H.; Wang, B. Empirical Analysis of Carbon Emission Accounting and Influencing Factors of Energy Consumption in China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2467. [\[CrossRef\]](#)
- Wang, S.; Zhao, T.; Zheng, H.; Hu, J. The STIRPAT Analysis on Carbon Emission in Chinese Cities: An Asymmetric Laplace Distribution Mixture Model. *Sustainability* **2017**, *9*, 2237. [\[CrossRef\]](#)
- McGee, J.A.; Clement, M.T.; Besek, J.F. The impacts of technology: A re-evaluation of the STIRPAT model. *Environ. Sociol.* **2015**, *1*, 81–91. [\[CrossRef\]](#)
- Han, J.; Du, T.; Zhang, C.; Qian, X. Correlation analysis of CO₂ emissions, material stocks and economic growth nexus: Evidence from Chinese provinces. *J. Clean. Prod.* **2018**, *180*, 395–406. [\[CrossRef\]](#)

18. Duro, J.A.; Padilla, E. International inequalities in per capita CO₂ emissions: A decomposition methodology by Kaya factors. *Energy Econ.* **2006**, *28*, 170–187. [[CrossRef](#)]
19. Hwang, Y.; Um, J.-S.; Schlüter, S. Evaluating the Mutual Relationship between IPAT/Kaya Identity Index and ODIAC-Based GOSAT Fossil-Fuel CO₂ Flux: Potential and Constraints in Utilizing Decomposed Variables. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5976. [[CrossRef](#)]
20. National Research Council. *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements*; The National Academies Press: Washington, DC, USA, 2010; p. 124.
21. Zaleski, P.; Yash, C. Circular Economy in Poland: Profitability Analysis for Two Methods of Waste Processing in Small Municipalities. *Energies* **2020**, *13*, 5166. [[CrossRef](#)]
22. Andrejiová, M.; Grincova, A.; Marasová, D. Study of the Percentage of Greenhouse Gas Emissions from Aviation in the EU-27 Countries by Applying Multiple-Criteria Statistical Methods. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3759. [[CrossRef](#)]
23. European Commission. *Communication from the Commission to the European Parliament, the European Council, the Council, The European Economic and Social Committee and the Committee of the Regions—The European Green Deal*; European Commission: Brussels, Belgium, 2019.
24. Sadik-Zada, E.R.; Gatto, A. Energy Security Pathways in South East Europe: Diversification of the Natural Gas Supplies, Energy Transition, and Energy Futures. In *From Economic to Energy Transition*; Springer: Berlin, Germany, 2020; pp. 491–514.
25. Mahony, T.O. Decomposition of Ireland's carbon emissions from 1990 to 2010: An extended Kaya identity. *Energy Policy* **2013**, *59*, 573–581. [[CrossRef](#)]
26. Wang, W.; Kuang, Y.; Huang, N.; Zhao, D. Empirical research on decoupling relationship between energy-related carbon emission and economic growth in Guangdong province based on extended Kaya identity. *Sci. World J.* **2014**, *2014*, 782750. [[CrossRef](#)]
27. Ziemele, J.; Gravelins, A.; Blumberga, D. Decomposition Analysis of District Heating System based on Complemented Kaya Identity. *Energy Procedia* **2015**, *75*, 1229–1234. [[CrossRef](#)]
28. Hwang, Y.; Um, J.-S. Performance evaluation of OCO-2 XCO₂ signatures in exploring casual relationship between CO₂ emission and land cover. *Spat. Inf. Res.* **2016**, *24*, 451–461. [[CrossRef](#)]
29. Van Amstel, A.; Olivier, J.; Janssen, L. Analysis of Differences between National Inventories and an Emissions Database for Global Atmospheric Research (EDGAR). *Environ. Sci. Policy* **1999**, *2*, 275–293. [[CrossRef](#)]
30. Hwang, Y.-S.; Lee, J.-J.; Park, S.-I.; Um, J.-S. Exploring Explainable Range of In-situ Portable CO₂ Sensor Signatures for Carbon Stock Estimated in Forestry Carbon Project. *Sens. Mater.* **2019**, *31*, 3773. [[CrossRef](#)]
31. Hwang, Y.; Um, J.-S. Comparative evaluation of OCO-2 XCO₂ signature between REDD+ project area and nearby leakage belt. *Spat. Inf. Res.* **2017**, *25*, 693–700. [[CrossRef](#)]
32. Li, Z.; Fu, J.; Lin, G.; Jiang, D.; Liu, K.; Wang, Y. Multi-Scenario Analysis of Energy Consumption and Carbon Emissions: The Case of Hebei Province in China. *Energies* **2019**, *12*, 624. [[CrossRef](#)]
33. Yokota, T.; Yoshida, Y.; Eguchi, N.; Ota, Y.; Tanaka, T.; Watanabe, H.; Maksyutov, S. Global Concentrations of CO₂ and CH₄ Retrieved from GOSAT: First Preliminary Results. *SOLA* **2009**, *5*, 160–163. [[CrossRef](#)]
34. Takagi, H.; Saeki, T.; Oda, T.; Saito, M.; Valsala, V.; Belikov, D.; Saito, R.; Yoshida, Y.; Morino, I.; Uchino, O.; et al. On the Benefit of GOSAT Observations to the Estimation of Regional CO₂ Fluxes. *SOLA* **2011**, *7*, 161–164. [[CrossRef](#)]
35. Maksyutov, S.; Takagi, H.; Valsala, V.K.; Saito, M.; Oda, T.; Saeki, T.; Belikov, D.A.; Saito, R.; Ito, A.; Yoshida, Y.; et al. Regional CO₂ flux estimates for 2009–2010 based on GOSAT and ground-based CO₂ observations. *Atmos. Chem. Phys.* **2013**, *13*, 9351–9373. [[CrossRef](#)]
36. Oda, T.; Maksyutov, S.; Andres, R.J. The Open-source Data Inventory for Anthropogenic CO₂, version 2016 (ODIAC2016): A global monthly fossil fuel CO₂ gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth Syst. Sci. Data* **2018**, *10*, 87–107. [[CrossRef](#)]
37. Oda, T.; Maksyutov, S. A very high-resolution (1 km × 1 km) global fossil fuel CO₂ emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.* **2011**, *11*, 543–556. [[CrossRef](#)]
38. Oda, T.; Bun, R.; Kinakh, V.; Topylko, P.; Halushchak, M.; Marland, G.; Lauvaux, T.; Jonas, M.; Maksyutov, S.; Nahorski, Z.; et al. Errors and uncertainties in a gridded carbon dioxide emissions inventory. *Mitig. Adapt. Strateg. Glob. Chang.* **2019**, *24*, 1007–1050. [[CrossRef](#)]

39. Hwang, Y.; Um, J.-S. Exploring causal relationship between landforms and ground level CO₂ in Dalseong forestry carbon project site of South Korea. *Spat. Inf. Res.* **2017**, *25*, 361–370. [CrossRef]
40. Park, S.-I.; Hwang, Y.; Um, J.-S. Utilizing OCO-2 satellite transect in comparing XCO₂ concentrations among administrative regions in Northeast Asia. *Spat. Inf. Res.* **2017**, *25*, 459–466. [CrossRef]
41. Hwang, Y.; Um, J.-S. Evaluating co-relationship between OCO-2 XCO₂ and in situ CO₂ measured with portable equipment in Seoul. *Spat. Inf. Res.* **2016**, *24*, 565–575. [CrossRef]
42. Park, A.-R.; Joo, S.-M.; Hwang, Y.; Um, J.-S. Evaluating seasonal CH₄ flow tracked by GOSAT in Northeast Asia. *Spat. Inf. Res.* **2018**, *26*, 295–304. [CrossRef]
43. GOSAT Project (GOSAT Data Archive Service). Available online: https://data2.gosat.nies.go.jp/index_en.html (accessed on 3 November 2020).
44. Kaya, Y.; Yokobori, K. *Environment, Energy, and Economy: Strategies for Sustainability*; United Nations University Press Tokyo: Tokyo, Japan, 1997.
45. Ma, M.; Cai, W. What drives the carbon mitigation in Chinese commercial building sector? Evidence from decomposing an extended Kaya identity. *Sci. Total Environ.* **2018**, *634*, 884–899. [CrossRef]
46. Raupach, M.R.; Marland, G.; Ciais, P.; Le Quéré, C.; Canadell, J.G.; Klepper, G.; Field, C.B. Global and regional drivers of accelerating CO₂ emissions. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10288. [CrossRef]
47. Raupach, M.R.; Canadell, J.G.; Le Quéré, C. Anthropogenic and biophysical contributions to increasing atmospheric CO₂ growth rate and airborne fraction. *Biogeosciences* **2008**, *5*, 1601–1613. [CrossRef]
48. Le Quéré, C.; Andrew, R.M.; Friedlingstein, P.; Sitch, S.; Pongratz, J.; Manning, A.C.; Korsbakken, J.I.; Peters, G.P.; Canadell, J.G.; Jackson, R.B.; et al. Global Carbon Budget 2017. *Earth Syst. Sci. Data* **2018**, *10*, 405–448. [CrossRef]
49. Peters, G.P.; Andrew, R.M.; Canadell, J.G.; Fuss, S.; Jackson, R.B.; Korsbakken, J.I.; Le Quéré, C.; Nakicenovic, N. Key indicators to track current progress and future ambition of the Paris Agreement. *Nat. Clim. Chang.* **2017**, *7*, 118–122. [CrossRef]
50. IEA. *Energy Statistics of OECD Countries 2015*; IEA: Paris, France, 2015.
51. IEA. *Energy Statistics of Non-OECD Countries 2015*; IEA: Paris, France, 2015.
52. MacKinnon, D.P.; Fairchild, A.J.; Fritz, M.S. Mediation analysis. *Annu. Rev. Psychol.* **2007**, *58*, 593–614. [CrossRef]
53. Namazi, M.; Namazi, N. Conceptual Analysis of Moderator and Mediator Variables in Business Research. *Procedia Econ. Financ.* **2016**, *36*, 540–554. [CrossRef]
54. MacKinnon, D.P. *Introduction to Statistical Mediation Analysis*; Routledge: Abingdon, UK, 2008.
55. Hayes, A.F. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*; Guilford Publications: New York, NY, USA, 2017.
56. Demming, C.; Jahn, S.; Boztug, Y. Conducting Mediation Analysis in Marketing Research. *Mark. ZFP* **2017**, *39*, 76–98. [CrossRef]
57. Preacher, K.J.; Hayes, A.F. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instrum. Comput.* **2004**, *36*, 717–731. [CrossRef] [PubMed]
58. Zhao, X.; Lynch, J.G.; Chen, Q.; Deighton, J. served as editor and Fitzsimons, G. served as associate editor for this article. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *J. Consum. Res.* **2010**, *37*, 197–206. [CrossRef]
59. Purcel, A.-A. New insights into the environmental Kuznets curve hypothesis in developing and transition economies: A literature survey. *Environ. Econ. Policy Stud.* **2020**, *22*, 585–631. [CrossRef]
60. Georgiev, E.; Mihaylov, E. Economic growth and the environment: Reassessing the environmental Kuznets Curve for air pollution emissions in OECD countries. *Lett. Spat. Resour. Sci.* **2015**, *8*, 29–47. [CrossRef]
61. Choi, E.; Heshmati, A.; Cho, Y. An empirical study of the relationships between CO₂ emissions, economic growth and openness. *J. Environ. Policy* **2010**, *10*. [CrossRef]
62. Robalino-López, A.; Mena-Nieto, Á.; García-Ramos, J.-E.; Golpe, A.A. Studying the relationship between economic growth, CO₂ emissions, and the environmental Kuznets curve in Venezuela (1980–2025). *Renew. Sustain. Energy Rev.* **2015**, *41*, 602–614. [CrossRef]
63. Graham, M.H. Confronting multicollinearity in ecological multiple regression. *Ecology* **2003**, *84*, 2809–2815. [CrossRef]
64. Sadik-Zada, E.R. Drivers of CO₂-Emissions in Fossil Fuel Abundant Settings: (Pooled) Mean Group and Nonparametric Panel Analyses. *Energies* **2020**, *13*, 3956. [CrossRef]

65. Wooldridge, J.M. *Introduction to Econometrics: Europe, Middle East and Africa Edition*; Cengage Learning: Boston, MA, USA, 2014.
66. Roos, M.W.M. Endogenous Economic Growth, Climate Change and Societal Values: A Conceptual Model. *Comput. Econ.* **2018**, *52*, 995–1028. [[CrossRef](#)]
67. Fischer-Kowalski, M.; Amann, C. Beyond IPAT and Kuznets Curves: Globalization as a Vital Factor in Analysing the Environmental Impact of Socio-Economic Metabolism. *Popul. Environ.* **2001**, *23*, 7–47. [[CrossRef](#)]
68. Ahi, P.; Searcy, C. A comparative literature analysis of definitions of green and sustainable supply chain management. *J. Clean. Prod.* **2013**, *52*, 329–341. [[CrossRef](#)]
69. Gong, R.; Xue, J.; Zhao, L.; Zolotova, O.; Ji, X.; Xu, Y. A Bibliometric Analysis of Green Supply Chain Management Based on the Web of Science (WOS) Platform. *Sustainability* **2019**, *11*, 3459. [[CrossRef](#)]
70. Zimon, D.; Tyan, J.; Sroufe, R. Implementing Sustainable Supply Chain Management: Reactive, Cooperative, and Dynamic Models. *Sustainability* **2019**, *11*, 7227. [[CrossRef](#)]
71. Sadik-Zada, E.R.; Gatto, A. The puzzle of greenhouse gas footprints of oil abundance. *Socio-Econ. Plan. Sci.* **2020**, 100936. [[CrossRef](#)]
72. Karl, H.; Ranné, O. European environmental policy between decentralisation and uniformity. *Intereconomics* **1997**, *32*, 159–169. [[CrossRef](#)]
73. Sadik-Zada, E.R.; Ferrari, M. Environmental Policy Stringency, Technical Progress and Pollution Haven Hypothesis. *Sustainability* **2020**, *12*, 3880. [[CrossRef](#)]
74. Scarpato, D.; Civero, G.; Rusciano, V.; Risitano, M. Sustainable strategies and corporate social responsibility in the Italian fisheries companies. *Corp. Soc. Responsib. Environ. Manag.* **2020**, *27*, 2983–2990. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).