Economic Factorial Analysis of CO$_2$ Emissions: The Divisia Index with Interconnected Factors Approach

Alexander Y. Vaninsky

Abstract—This paper presents a method of economic factorial analysis of the CO$_2$ emissions based on the extension of the Divisia index to interconnected factors. This approach, contrary to the Kaya identity, considers three main factors of the CO$_2$ emissions: gross domestic product, energy consumption, and population - as equally important, and allows for accounting of all of them simultaneously. The three factors are included into analysis together with their carbon intensities that allows for obtaining a comprehensive picture of the change in the CO$_2$ emissions. A computer program in R-language that is available for free download serves automation of the calculations. A case study of the U.S. carbon dioxide emissions is used as an example.

Keywords—CO$_2$ emissions, Economic analysis, Factorial analysis, Divisia index, Interconnected factors.

I. INTRODUCTION

Environmental degradation, global warming, and climate change, which economic development partially stimulates, may heavily impact wellbeing. For example, publication [11] states that people's average life expectancy has been reduced by five years in northern China due to heavy air pollution. The latest figures state that if the rise in global average temperature exceeds 4°C, irreversible catastrophic consequences on a global scale may be imminent. To achieve a reasonable sustainable balance between economic development and environment protection, the detailed analysis of the CO$_2$ and other gas emissions is necessary.

Factorial economic analysis is a widely used tool for finding the main factors of the gas emissions. Since the CO$_2$ emissions constitute the greatest part, the gas emissions are usually expressed as CO$_2$- equivalent. Using the Kaya identity [5] is a widespread approach, [4], [7]. It expresses total carbon dioxide emissions as a product of carbon intensity of energy (CO$_2$/E), the energy intensity of economic activity (E/GDP), GDP per capita (GDP/P), and population (P):

\[ CO2 = (CO2/E) \times (E/GDP) \times (GDP/P) \times P \]

and computes the contributions of the factors to the change in the CO$_2$ emissions.

The Kaya-identity approach may be criticized from different perspectives. First, only population is included as a quantitative indicator. Neither energy, nor GDP is considered in the framework of the factorial model (1). As a result, if the amount of energy $E$ increases, other things equal, the model does not reveal the increase in the CO$_2$. It just decreases the first term (CO$_2$/E) and increases the second (E/GDP) that is counterintuitive. Second, different factor models similar to that given by (1) may be suggested leading to different factorial decomposition.

It may be noted also that while many publications in the field of CO$_2$ emissions cite the publication [5], the methodology may be traced back to the foundations of the index number theory developed in [6], [10], see [2] for review. The factor models similar to (1) have also been studied extensively by A.D. Sheremet and his school of economic analysis [13].

Publication of Divisia [3] suggested consideration of the factors as functions of time $t$, and used the formula of the complete differential as the base of factorial analysis. Let $Z$- an indicator under analysis (resulting indicator), be a product of the factors $X_1, X_2, \ldots, X_n$:

\[ Z = X_1X_2\ldots X_n, \]

then

\[ \Delta Z = Z_1 - Z_0 = \int dZ = \int \frac{X_1X_2\ldots X_n}{L} dX_1 + \int \frac{X_1X_2\ldots X_n}{L} dX_2 + \ldots + \int \frac{X_1X_2\ldots X_n}{L} dX_n \]

where $L$ is a curve of the factors' change, with each additive term related to a factor $X_i$:

\[ \Delta Z[X_i] = \int \frac{X_1X_2\ldots X_i\ldots X_n}{L} dX_i \] (4)

The factorial decomposition suggested in [3] is

\[ \Delta Z = \Delta Z[X_1] + \Delta Z[X_2] + \ldots + \Delta Z[X_n] \]

This approach has been developed further in [14] to include arbitrary continuously differential functions.

It was shown, however, in publication of Meirovich [9] that the Laspeyres-Paasche approach leads to the results
contradicting economic common sense when structural change is analyzed. An empirical formula for computations has been suggested that corrected the mentioned paradox. In publication [18] this empirical formula has been extended to a general case of interconnected factors. Mathematical theory was developed in [16], [17]; its brief description is given in [8]. We refer to it as a generalized Divisia index method (GDIM) with interconnected factors in this paper below. We show how the GDIM-based approach suggested in [15] for factorial analysis of production may be adapted to the analysis of CO₂ emissions to include all three factors: GDP, energy consumptions and population and their CO₂ intensities.

The paper is organized as follows. Section II provides the description of the GDIM, Section III applies it to the analysis of CO₂ emissions, and Section IV offers an example of analysis of the U.S. carbon dioxide emissions. A computer program in R-language is given in the Appendix.

II. GENERALIZATION OF THE DIVISIA INDEX METHOD TO INTERCONNECTED FACTORS


\[ Z = f(X) = f(X_1, ..., X_n) \]  

and received the following factorial decomposition:

\[ \Delta Z = Z_1 - Z_0 = \left[ \int f'_1 dX_1 + \int f'_2 dX_2 + ... + \int f'_n dX_n \right] \]  

where \( f'_i \) is a partial derivative with respect to the \( i \)-th argument. Formula (7) may be rewritten as

\[ \Delta Z = \nabla Z' \cdot dX \]  

where \( \Delta Z \) is a row vector with coordinates

\[ \Delta Z(X_i) = \int f'_i dX_i \]  

and

\[ \nabla Z = < f_1, ..., f_n > \]  

is a gradient vector of the function \( f(X_1, ..., X_n) \); upper index \( T \) stands for the transposition, and \( dX \) is a diagonal matrix with elements \( dX_1, dX_2, ..., dX_n \).

It was shown in [18] that the paradox mentioned in [9] is not resolved in the framework of factorial decomposition (8)-(10) as well, and it was suggested that the reason for the paradox is the interconnection of the structural factors by an equation

\[ X_1 + ... + X_n = 1 \]  

It was proposed in [18] to extend the statement of the problem (6)-(7) with the equations of the factors' interconnection

\[ \Phi_j(X_1, ..., X_n) = 0, j = 1, ..., k \]  

or in a vector form,

\[ \Phi(X) = 0 \]  

It was hypothesized, at the intuitive level, that to avoid the paradox, the vector of the infinitesimal factors' change

\[ dX = < dX_1, ..., dX_n > \]  

should be projected on the surface defined by the equations of the factors' interconnection (12)-(13). As a result, the following formula was obtained:

\[ \Delta Z / X | \Phi = \left[ \nabla Z' (I - \Phi_x \Phi_x^+) \right] dX \]  

where the row vector \( \Delta Z / X | \Phi \) stands for the factorial decomposition of the change in the resulting indicator \( Z \) in the presence of the factors' interconnections (13), \( \Phi_x \) is a Jacobian matrix for the matrix \( \Phi(X) \):

\[ (\Phi_x)_i^j = \frac{\partial \Phi_j}{\partial X_i} \]  

upper index “\( ^+ \)” denotes the generalized inverse matrix, and \( I \) is an identity matrix. If the columns of the matrix \( \Phi_x \) are linearly independent, then

\[ \Phi_x^+ = (\Phi_x^T \Phi_x)^{-1} \Phi_x^T \]  

see [1] for detail.

Axiomatic theory of the factorial decomposition in the presence of the factors' interconnection was developed in [16], [17]; its brief description may be found in [8]. Since the (15) uses an operator of projection on a surface, the factors should be measured in relative units, see [17] for detail.

Publication [15] applies the suggested approach to factorial decomposition of the change in production by the factors of fixed and variable assets, and labor. In this paper we accommodate it to analysis of CO₂ emissions by the factors of GDP, energy, and population.

III. FACTORIAL DECOMPOSITION OF THE CHANGE IN THE CO₂ EMISSIONS

In this section we apply the approach suggested in [15] to economic analysis of the CO₂ emissions. We begin with an
observation that CO₂ emissions may be presented in either of the three ways:

\[
\text{CO}_2 = \frac{\text{CO}_2}{\text{GDP}} \cdot \text{GDP} = \frac{\text{CO}_2}{\text{Energy}} \cdot \text{Energy} = \frac{\text{CO}_2}{\text{Population}} \cdot \text{Population}
\]

\[ \text{(18)} \]

Our objective is to incorporate all of them in factorial analysis in a symmetric manner.

For readability, we use in the following transformations the following denotations: \( Z = \text{CO}_2 \), \( X_1 = \text{GDP} \), \( X_4 = \text{Energy} \) consumption, \( X_6 = \text{Population} \), and \( X_2, X_5 \) - their correspondent carbonization intensities: \( X_2 = \frac{\text{CO}_2}{\text{GDP}} \), \( X_4 = \frac{\text{CO}_2}{\text{Energy}} \), and \( X_5 = \frac{\text{CO}_2}{\text{Population}} \). In terms of the newly defined variables, (18) becomes

\[
Z = X_1 X_2 = X_3 X_4 = X_5 X_6.
\]

(19)

To apply the GDIM, we separate these equations into a factor model and equations of the factors’ interconnections as follows:

\[
Z = X_1 X_2, \quad X_1 X_2 = X_3 X_4 = X_5 X_6.
\]

(20)

As shown in [15], any pair of factors may be chosen in the first equation in (20) without change in the final result. The last two equations in (20) form a system of (12) that may also be written in a vector form (13).

In terms of the variables \( X_i \) the gradient of the function \( Z(X) \) and the Jacobian matrix \( \Phi_i \) take the form

\[
\nabla Z = \begin{bmatrix} X_2, X_1, 0, 0, 0, 0 \end{bmatrix} \thinspace >^T
\]

\[ \text{(21)} \]

\[
\Phi_i = \begin{bmatrix} X_2, X_1, -X_4, -X_2, 0, 0, 0, -X_6, -X_5 \end{bmatrix}
\]

(22)

The factorial decomposition given by (15) becomes

\[
\Delta Z [X_i | \Phi_i] = \int_L (X_1 \cdot B_{21} + X_2 B_{11}) dX_i
\]

(23)

where \( \Delta Z [X_i | \Phi_i] \) is the \( i \)-th coordinate of the vector of factorial decomposition (15), \( i = 1, \ldots, 6 \), and \( B_{11} \) and \( B_{21} \) are as follows:

\[
\begin{align*}
B_{11} &= \left( X_1 \left( X_3 + X_4 + X_6 \right) - \left( X_3 + X_4 \right) \left( X_3 + X_6 \right) \right) \\
B_{12} &= \left( -X_1 X_2 \left( X_3 + X_4 + X_6 \right) + \left( X_3 + X_4 \right) \left( X_3 + X_6 \right) \right) \\
B_{13} &= \left( X_1 X_2 \left( X_3 + X_4 \right) \right) \\
B_{14} &= \left( X_1 X_2 \left( X_3 + X_6 \right) \right) \\
B_{15} &= \left( X_1 X_2 \left( X_4 + X_6 \right) \right) \\
B_{16} &= \left( X_1 X_2 \left( X_4 + X_4 \right) \right)
\end{align*}
\]

(24)

To make calculations, we need to parameterize the curve of the factors’ dynamics \( L \). It is typical for the analytical purposes to assume a linear or an exponential change in the quantitative indicators in time. In this paper below we use a model time \( t \) that varies in the interval \([0,1]\). It may be shown that the length of the time-interval does not affect the final result. In case of the linear dynamics,

\[
\begin{align*}
Q(t) &= Q(0) + (Q(1) - Q(0)) t, \\
\frac{dQ(t)}{dt} &= Q(1) - Q(0),
\end{align*}
\]

(25)

for the exponential change in time,

\[
\begin{align*}
Q(t) &= Q(0) \cdot (Q(1)/Q(0))^t, \\
\frac{dQ(t)}{dt} &= \ln(Q(1)/Q(0)) \cdot (Q(1)/Q(0))^t, \\
\end{align*}
\]

(26)

where \( Q \) stands for the \text{CO}₂, GDP, energy, or population, as appropriate. The dynamics of the carbonization intensities are obtained as a ratio of \text{CO}_2 to the corresponding factor. For example, the change in time of the GDP carbonization indicator (\text{CO}_2/GDP) is this

\[
(\text{CO}_2/\text{GDP})(t) = \text{CO}_2(t)/\text{GDP}(t)
\]

(27)

If the carbonization intensity does not follow the exponential change in time, a formula for the differential of the ratio of two functions should be used in (23).

The factorial decomposition is obtained by the substitution of the \( t \)-parameterization (25) - (27) into (23). The result is obtained in terms of the relative change in the base value of \text{CO}_2. A computer program in R-language [12] given in the Appendix performs calculations for the exponential change in \text{CO}_2, GDP, energy, and populations. It may be adjusted for any other dynamics of the quantitative indicators.

IV. THE U.S. CASE STUDY

In this section we apply the suggested approach to the analysis of the \text{CO}_2 emissions in the United States in 2010 as compared to the 1990. We investigate to what extent the change in GDP, energy consumption, population, and their carbonizations affected the increase in the \text{CO}_2 emissions for
this period. The obtained results are aimed to serve both analytical goals and the objectives of the environmental protection policy.

Statistical data were collected from the websites of the U.S. Energy Information Administration [20] and the U.S. Bureau of Economic Analysis [19]. They are shown in columns 2 and 3 of the Table I, rows 1 through 4. Carbon intensity indicators were calculated as the ratios of the CO2 emissions to the corresponding quantitative indicators: CO2/GDP, CO2/Energy, and CO2/Population, respectively. They are shown in the rows 5 - 8, of the columns 2 and 3, correspondingly

<table>
<thead>
<tr>
<th>Factors(*)</th>
<th>1990 (Base)</th>
<th>2010</th>
<th>Change</th>
<th>Change, %</th>
<th>Contribution to the change in CO2(^2)</th>
<th>% of the base value</th>
<th>min metric tons</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td></td>
</tr>
<tr>
<td>CO2</td>
<td>5047.1</td>
<td>5633.6</td>
<td>586.5</td>
<td>11.6</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>8027.1</td>
<td>13063.0</td>
<td>5035.9</td>
<td>62.7</td>
<td>16.0</td>
<td>806.1</td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>84.5</td>
<td>98.4</td>
<td>13.9</td>
<td>16.4</td>
<td>5.6</td>
<td>281.1</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>249.6</td>
<td>310.1</td>
<td>60.5</td>
<td>24.2</td>
<td>7.9</td>
<td>397.0</td>
<td></td>
</tr>
<tr>
<td>CO2/GDP</td>
<td>0.6288</td>
<td>0.4313</td>
<td>-0.1975</td>
<td>-31.4</td>
<td>-12.4</td>
<td>-624.1</td>
<td></td>
</tr>
<tr>
<td>CO2/Energy</td>
<td>59.7361</td>
<td>57.2811</td>
<td>-2.4549</td>
<td>-4.1</td>
<td>-1.5</td>
<td>-77.7</td>
<td></td>
</tr>
<tr>
<td>CO2/Population</td>
<td>20.2208</td>
<td>18.1670</td>
<td>-2.0537</td>
<td>-10.2</td>
<td>-3.9</td>
<td>-195.9</td>
<td></td>
</tr>
<tr>
<td>Total(^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.6</td>
<td>586.5</td>
<td></td>
</tr>
<tr>
<td>Notes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 1\(^*\) Units of measurement: CO2 - million metric tons, GDP - billion $2005, Energy - quadrillion Btu, Population - million people, carbon intensities - correspondingly, as appropriate.  
2 Column 6 is calculated by using the algorithm provided in the text and the computer program given in the Appendix section. Column 7 is obtained by multiplication of the column 6 by the level of the CO2 emissions in the base year.  
3 Total is equal to the change or the percentage change in the CO2 emissions, correspondingly.

Columns 4 and 5 comprise change and percentage change in the corresponding indicators. Column 6 was calculated by using the algorithm provided in the text and computer program given in the Appendix section. Its entries show percentages of the relative change in the CO2 emissions that are due to the impact of particular factors. Column 7 contains the corresponding amounts of the CO2 emissions. They were obtained as the products of the corresponding entries in column 6 by the base value of the CO2 emissions of 5047.1 million metric tons.

An example of calculations is as follows. The CO2 emissions increased in 2010 as compared to the 1990 by 586.5 million metric tons. This amount constituted 11.6% of the base value, see row 1 of the table 1. Due to the increase in GDP by 13063.0 - 8027.1 = 5035.9 billion $2005 the amount of the CO2 emissions has increased by 16.0% of the base value, as calculated by the R-program. This is equal to 16% of 5047.1 million metric tons or 806.1 million metric tons, as shown in columns 6 and 7, row 2, respectively. The CO2 intensity of the GDP has decreased from 0.6288 to 0.4313 million metric tons/billion $2005. This has led to the decrease in the CO2 emissions by 12.4% of the base value, calculated by the R-program, or 12.4% 5047.1 = 624.1 million metric tons of CO2, as shown in row 5 columns 6 and 7 of the Table I. Calculations for the energy and population indicators were conducted in the same way. The totals of all factorial components are equal to the change in the CO2 emissions: 11.6% of the base value or 586.5 million metric tons.

In this example, all quantitative indicators have increased leading to the increase in the CO2 missions. However, the percentage increase in the factors had not resulted in the equal percentage increase in the CO2 emissions. Thus, the increase by 62.7% in the GDP resulted in just 16.0% increase in the CO2 emissions.

All carbon intensity indicators have decreased. For example, the carbon intensity of the GDP has decreased by 31.4% that has led to the decrease in the CO2 emissions by 12.4% or 624.1 million metric tons of CO2.

In general, it may be mentioned that the direction of the factor's impact is the same as the change in the factor itself: an increase in a factor leads to the increase in the CO2 emissions and vice versa. But quantitatively, the change in the CO2 emissions is smaller. The latter observation is due to two reasons. First, a part of the factor's impact on the CO2 emissions results from the change in the paring indicator. Second, the interconnections among the groups of the factors - GDP, energy, or population, respectively - result in the transfer of the part of the change in each group to other groups, rather than affecting the CO2 emissions.

The international energy outlook report of the U.S. Energy Information Administration for 2011 [4] states that the main focus of policymakers is decreasing the energy intensity of economic output (Energy/GDP) and carbon dioxide intensity of the energy (CO2/E). The proposed approach provides a wider spectrum of opportunities and suggests estimations of the boundaries for each option. As follows from the Table I, the main factors of the change in the CO2 emissions for the period of 1990 - 2010 were growth of GDP (16.0%), population (7.9%), and energy (5.6%), combined with the decrease in the carbonization of GDP (CO2/GDP, -12.4%). Since governments pursue the goal of GDP growth and leave the increase in population unaffected, this observation shows that the main factors of the decrease in the CO2 emissions are energy saving and decarbonization of the GDP. If the main tendencies are held constant during the next 20 years, their potential, theoretically, is 5.6% + 12.4% = 18.0% of the 2010 CO2 emissions level. Practically, this means that the energy policy should be directed to energy saving and cleaner technologies of production. It may be noted also that a potential impact of energy decarbonization (CO2/Energy) that is usually mentioned as a key factor, was found in our research to be relatively small - just 1.5% of the base level. These results, if confirmed by the subsequent research, may lead to the change in the priorities of the U.S. environmental protection policy.

V. CONCLUSIONS

This paper presents a tool of factorial decomposition of the
CO₂ emissions by interconnected factors: GDP, energy consumption, population, and their carbon intensities. Mathematically, the suggested approach is an extension of the Divisia index to the interconnected factors. A computer program in the R-language provided in the Appendix makes the computations. A case study of the U.S. carbon dioxide emissions in 2010 versus 1990 illustrates the details of the calculations and offers an example of the analysis.

APPENDIX

This section presents a program in the R language [12], version 2.15.0, that performs factorial decomposition of the change in the CO₂ emissions. The program assumes exponential dynamics of all quantitative indicators: CO₂ emissions, GDP, energy consumption, and population. The results are obtained in terms of the relative change in the base value of CO₂. In this version of the program, the statistical data were included in the R script file. In general, they should be imported from a spreadsheet or database using the R-language tools. Symbol “#” stands for a comment line or a part of it.

# ------Input to the program --------------------------------------------
# ----------------------------------------------------------------------
# yearB <- 1990 # Base year
year <- 2010 # Year of calculations
C0 <- 5047.1 # CO₂ emissions, base year
C1 <- 5633.6 # CO₂ emissions, calculations' year
G0 <- 8027.1 # GDP, base year
G1 <- 13063.0 # GDP, calculations' year
E0 <- 84.49 # Energy, base year
E1 <- 98.35 # Energy, calculations' year
P0 <- 249.6 # Population, base year
P1 <- 310.1 # Population, calculations' year
# ----------------------------------------------------------------------
# --- Data in terms of the base year -------------------------------------
z1 <- C1/C0 # CO₂ emissions in terms of the base year
x1 <- G1/G0 # GDP in terms of the base year
x3 <- E1/E0 # Energy in terms of the base year
x5 <- P1/P0 # Population in terms of the base year
# Print out results and control numbers -----------------------------
# --- Control number should be equal to the rate of change of the CO2 emissions
yearB; year; "GDP"; Dz1; "CO2/GDP";Dz2; "Energy";
Dz3; "CO2/Energy"; Dz4; "Population"; Dz5; "CO2/Population"; Dz6;
"Check sum"; Dz1+Dz2+Dz3+Dz4+Dz5+Dz6; "Rate of change in CO2"; z1-1

# --- Print out results and control numbers -----------------------------
# "Contributions to the rate of change in CO2 emissions"
yearB; year; "GDP"; Dz1; "CO2/GDP"; Dz2; "Energy";
Dz3; "CO2/Energy"; Dz4; "Population"; Dz5; "CO2/Population"; Dz6;
"Check sum"; Dz1+Dz2+Dz3+Dz4+Dz5+Dz6; "Rate of change in CO2"; z1-1

# --- Print out results and control numbers -----------------------------
# "Contributions to the change in the CO2 emissions"
yearB; year; "GDP"; Dz1; "CO2/GDP"; Dz2; "Energy";
Dz3; "CO2/Energy"; Dz4; "Population"; Dz5; "CO2/Population"; Dz6;
"Check sum"; Dz1+Dz2+Dz3+Dz4+Dz5+Dz6; "Rate of change in CO2"; z1-1

REFERENCES


