Economics World, Oct.-Dec. 2024, Vol. 11, No. 4, 186-202
doi: 10.17265/2328-7144/2024.04.003

Distribution of Expenditures for Research and Development—Dynamic Analysis Based on Composite Concentration Coefficients

Jasmin Komić
University of Banja Luka, Banja Luka, Bosnia and Herzegovina

The dynamic analysis of distribution of expenditures for research and development aims to identify concentration tendencies in this domain. The analysis is performed by calculating composite concentration coefficients, as these coefficients allow simultaneous (synthesized) expression of existing concentrations in cases where multiple criteria for expressing concentration exist. In this work, the observation was made on a respectable sample of 52 countries in the period from 1997 to 2021. The following variables are included in the analysis: population, gross domestic product, gross domestic product per capita, gross domestic expenditure on research and development absolutely and per capita and human development index. Composite concentration coefficients were derived based on previously calculated concentration coefficients according to the criteria of population size and gross domestic product, gross domestic product per inhabitant and social development index. Calculation of composite concentration coefficients was performed by extrapolation based on standardized regression coefficients from the corresponding regression models. Based on the conducted analysis and obtained results, it can be concluded that a relatively high concentration of gross expenditures for research and development is evident, which has been significantly reduced in the observed time period.

Keywords: research, development, composite concentration coefficients

Introduction

During each analysis of expenditures on research and development, the fundamental and indisputable fact is that scientific research activity is one of the primary and most important drivers of growth and development, aiming to meet various societal needs, to raise the level of economic activity and the standard of living of the population (Romer, 1990; Grossman, & Helpman, 1991; Aghion, & Howitt, 1992; Lucas, 1988). Even under very conservative assumptions, it is difficult to find an average return below $4 per $1 spent on research and development. Accounting for health benefits, inflation bias, or international spillovers can bring the social returns to over $20 per $1 spent (Jones, & Summers, 2020; Hall, & Mairesse, 2009).

Regardless of variations in the rate of return, it is evident that research and development is an important driver of economic growth Top of Form (Nekrep, Strašek, & Boršič, 2018). To track funds allocated to research

---

Jasmin Komić, Ph.D., professor, Faculty of Economics, University of Banja Luka, Banja Luka, Bosnia and Herzegovina.
Correspondence concerning this article should be addressed to Jasmin Komić, 80 Braće Jugovića str., 78000 Banja Luka, Bosnia and Herzegovin.

and development according to socioeconomic goals, the Nomenclature for the Analysis and Comparison of Scientific Programs and Budgets, as prescribed in the Frascati Manual (2015) is utilized.

Research and development encompass three types of activities: basic research, applied research, and experimental development (developmental research).

Research and development activities take place in four sectors: the business sector, the higher education sector, the government sector, and the non-profit sector.

The most important indicator in this regard shows the share of expenditure on research and development (R&D) in gross domestic product.

The concentration of gross expenditure on R&D is just one of many aspects of its observation. In the context mentioned above, this analysis points to one aspect of understanding the state and trends regarding expenditure on R&D and developmental trends in this regard on global and individual countries scales.

Methodology

The Variables Included

The dynamic analysis of the distribution of R&D allocation is a complex issue and can be conducted in several ways depending on the research objective, the set of the included variables, the number and arrangement of the observed countries, the length of the observed period, etc. In this paper, the analysis was performed on the basis of the following variables: population (POPULATION), gross domestic product (GDP), gross domestic product per capita (GDP/cap), gross domestic expenditure on research and development (GERD) and social development index (HDI).

Gross Domestic Expenditure on Research and Development

Research and experimental development (R&D) involves systematic creative work undertaken to increase knowledge stocks, including knowledge about humanity, culture, and society, as well as new applications of existing knowledge (Frascati Manual, 2002) and its utilization to create new one.

Gross domestic expenditure on research and experimental development (Frascati Manual, 2002) (GERD) is an aggregate representing the total intramural expenditure on research and development carried out within the national territory during the observed period. It includes research and development conducted within the country and financed from abroad, but excludes payments for research and development conducted abroad.

Population

Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates.

Gross Domestic Product

The most commonly used macroeconomic aggregate in the System of National Accounts is Gross Domestic Product (GDP), which represents a measure of final production. Gross Domestic Product per capita (GDP/cap) is a relative indicator obtained by dividing the value of gross domestic product by the population. This provides a comparative indicator for comparing different countries.

Human Development Index

Human Development Index (HDI) (Human Development Reports, 1990; Stanton, 2007) is the result of the search for a common measure of economic and social development. It is an effort to quantify the entire socioeconomic aspect of progress, which refers to the achievements of the country in terms of the basic dimensions of social development (Janjić, 2013).

HDI includes three basic dimensions of social development, which relate to the opportunities people expect to achieve. These are health, education and standard of living. HDI is the geometric mean of the normalized indicators for each of the three dimensions.

Observation Period

The dynamic analysis was conducted based on time series of gross expenditure on R&D for the included countries in the period from 1997 to 2021. In every analysis of this type, the aim is to have time series as long as possible. Limitations in forming these time series primarily relate to the existence and availability of data.

Countries Included

Forming appropriate sets of valid, high-quality, up-to-date, and transparent data is particularly important and challenging in every research phase, as all methodological procedures and subsequent analyses cannot correct errors made during this phase. The basis of quantitative research in this study is the available data.

When forming appropriate databases for the observed time period, some countries lack data. In such cases, the author considered the criterion that countries with at most two non-adjacent missing data points be included in the analysis. For adjacent data points given as relative values, imputation was conducted using interpolation based on the geometric mean. Additionally, extrapolation based on a linear trend was performed for missing data at the beginning (8) and end (9) of the observed period. This way, with 19 data points (1.5% of the total number of data points) obtained in the previously described manner, an additional 14 countries were included in the analysis (26.9% of the total number of observed countries), which completely justifies this approach. Consequently, the total number of included countries in the analysis is 52.

The countries included in the analysis are: Argentina, Austria, Azerbaijan, Belgium, Bulgaria, Belarus, Czech Republic, Denmark, Estonia, Finland, France, Netherlands, Hong Kong, Croatia, India, Ireland, Italy, Israel, Japan, Armenia, South Korea, Canada, Kazakhstan, China, Cyprus, Kyrgyzstan, Colombia, Cuba, Kuwait, Latvia, Lithuania, Hungary, Mexico, Mongolia, Norway, Germany, Panama, Poland, Portugal, Romania, Russian Federation, USA, Singapore, North Macedonia, Serbia, Slovakia, Slovenia, Spain, Trinidad and Tobago, Turkey, Ukraine, and United Kingdom.

The observed set of countries has a very respectable coverage. In 2021, 56.6% of the world's population and 82.6% of the world's gross domestic product are covered.

This leads to the conclusion that the obtained results are valid and that it is possible to generalize the obtained results at the level of the entire set of countries worldwide. On the other hand, considering the different coverage regarding population and gross domestic product, there is a clear need for the application of composite concentration coefficients when analyzing concentration in terms of allocation for R&D.

Methods

During the observed period from 1997 to 2021, a dynamic analysis of the concentration of gross expenditure on R&D was conducted based on the Gini coefficient of concentration (Gini, 1921), as one of the most frequently used measures of concentration.

The Gini coefficient of concentration is a relative analytical measure used to express the present disparity or concentration in terms of degree and direction in the distribution or observed series.

To define the Gini index of concentration (Šošić, 2004), it is necessary to define the Lorenz curve (Lorenz, 1905), which is obtained by connecting points with coordinates:

$$(0,0),\left(\frac{1}{n},\frac{x_1}{\sum_{i=1}^{n}x_i}\right),\ldots,\left(\frac{i}{n},\frac{\sum_{j=1}^{i}x_j}{\sum_{i=1}^{n}x_i}\right),\ldots,(1,1)\text{ ; } i = 1, 2, \ldots, n.$$

Values for which the concentration coefficient is calculated should be sorted by size:

$$x_1 \le x_2 \le \ldots \le x_{i-1} \le x_i \le \ldots \le x_N\text{ ; } \; x_i \ge 0, \forall i\text{ , } \; \sum_{i=1}^{n} x_i > 0$$

Based on the above, the expressions for calculating the concentration coefficient for a series where the number of modalities and corresponding values are given as follows:

$$G = 1 - \sum_{i=1}^{n} p_i \left[F_T(T_i) + F_T(T_{i-1})\right], \quad F_T(T_0) = 0$$

$$G = \frac{2\sum_{i=1}^{n} i x_i - (n+1)\sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i}.$$

In the first expression $p_i = 1/n$, and $F_T(T_i)$ are the corresponding values from the cumulative proportion of subtotals.

If the numerical data relate to some quantitative continuous statistical characteristic (e.g., height, weight, etc.) and if they belong to the interval [a, b], then they cannot be ordered as an (infinite) sequence, which means that the corresponding Lorenz curve represents a real function *f(x)* on the interval [0, 1], so that concentration can be expressed by the expression (Kovačić, Opačić, & Marohnić):

$$G = 1 - 2\int_0^1 f(x)dx$$

The function $f(x)$ must be monotonically increasing and should satisfy the following properties:

$$f(0) = 0$$

$$f(1) = 1$$

$$0 \le f(x) \le x, \forall x \in [0,1]$$

and that it is Riemann/Lebesgue integrable on the interval [0, 1].

The Gini concentration coefficient can be calculated in several ways, one of which is based on the triangle method and is obtained from the expression (Somun-Kapetanović, 2008):

$$G = \frac{\sum p_i q_{i+1} - \sum q_i p_{i+1}}{10000}$$

where p_i and q_i denote the relative values from corresponding cumulants of the observed series.

The concentration coefficient takes values in the range from -1 to +1. The sign indicates concavity or convexity in the graphical representation of the distribution of obtained values. In the case of a symmetric distribution, this coefficient takes the value zero. When the concentration is maximum, the coefficient takes the value +1 or -1. Therefore, the more pronounced the disparity, the closer the value of this coefficient is to one.

The average annual growth rate was used as a measure of concentration changes in the observed time period (Lovrić, Komić, & Stević, 2017).

Simple linear correlation coefficient (Ibid) was used to measure the agreement of the variations of the obtained concentration coefficients and the variations of the included variables.

The geometric mean was used to obtain synthesized composite concentration coefficients (Ibid).

Composite Concentration Coefficients

Composite indices of concentration are calculated on the basis of concentration coefficients obtained on the basis of individual criteria (in this case, population size, gross domestic product and Human Development Index).

Composite concentration coefficients can be obtained as an average value, i.e. as an arithmetic or geometric mean, of the concentration coefficients according to the two previous criteria. This approach is relatively simple in computational terms, but it has a major drawback. It is reflected in the fact that the concentration coefficients according to one and the other criterion have the same weighting, that is, the same specific weight. This cannot be accepted as a correct attitude and therefore it is necessary to somehow give each series of concentration coefficients an appropriate weight, that is, a specific weight.

This can be done by recognizing the relationships of the variables involved and quantifying those relationships. As already stated, both the size of the population and the gross domestic product affect the allocation for research and development. For this reason, it is logical to form appropriate regression models that should quantify the regularity of these relationships and to obtain composite concentration coefficients based on the estimated parameters of the regression models by extrapolation of these models, i.e. as a linear combination of concentration coefficients based on individual criteria derived on the basis of standardized coefficients estimated regression models.

Composite coefficients of concentration were obtained based on estimated regression model (Lovrić, Komić, & Stević, 2017):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i,\ i = 1, 2, \ldots, n$$

where is:

Y_i: i-th dependent variable

$\beta_0, \beta_1, \beta_2, \ldots, \beta_k$: model parameters

ε_i: stochastic term or random error

k: number of explanatory variables.

Unlike the original regression coefficients, which are named by the unit of measure of the variables included in the model, the standardized regression coefficients are unnamed, that is, they are invariant quantities. These coefficients show how many standard deviations the dependent variable will change when the standard deviation of the predictor variable increases.

Coefficient standardization is usually done to answer the question of which of the independent variables has a greater effect on the dependent variable in a multiple regression analysis where the variables are measured in different units of measurement. It can also be considered a general measure of effect size, quantifying the “size” of the effect of one variable on another. For simple linear regression with orthogonal predictors, the standardized regression coefficient is equal to the correlation between the independent and dependent variables.

Standardized regression coefficients (Menard, 2004) are obtained based on the regression coefficients from the estimated regression model based on the expression:

$$\beta^* = \beta \frac{sd_X}{sd_Y}$$

in which *β* denotes the original regression coefficients, and *β** the standardized regression coefficients. Standard deviations of dependent (*Y*) and independent (*X*) variables are marked with *sd*.

The weighting of concentration coefficients based on standardized regression coefficients is performed by extrapolation of values based on the estimated regression equation, that is by calculating composite concentration coefficients as a linear combination of previously obtained concentration coefficients based on individual criteria, and based on the value of standardized regression coefficients.

The problems that can arise during the formation and evaluation of the regression model are multiple. This primarily refers to its statistical significance, specification, occurrence of multicollinearity and heteroskedasticity.

If the analysis of variance shows that the obtained model is not statistically significant, then it is not possible for the evaluated model to be used to obtain composite concentration coefficients. In that case, the composite concentration coefficients should be calculated as average values, primarily using the geometric mean.

In this type of analysis, the starting point is the inclusion of criterion variables in the model and its linearity, which somewhat relaxes the mentioned problems.

Multicollinearity represents the mutual linear dependence of the explanatory variables in a linear regression model and the lack of independent variation in the explanatory variables to separate their separate effects on the dependent variable. To test the presence of multicollinearity, the Farrar and Glauber test is most often used (Farrar, & Glauber, 1967). Method of principal components (Jovičić, 2002) enables the detection and elimination of undesirable consequences of multicollinearity. To diagnose the presence and level of multicollinearity in practice, Tolerance and VIF are most often used (Field, 2018). Tolerance shows how much of the variance of the observed explanatory variable is not explained by the variances of other independent variables in the model and is calculated according to the formula $1\text{-}R_i^2$ for each variable. If this value is very small (less than 0.10), it indicates a very high correlation of that variable with other independent variables, that is, the presence of harmful effects of multicollinearity. VIF (Variance Inflation Factor) represents the variance increase factor and is obtained as the reciprocal of Tolerance, i.e. $1/(1\text{-}R_i^2)$. If VIF values are greater than 10, this indicates the presence of harmful multicollinearity.

Heteroskedasticity (Jovičić, & Dragutinović-Mitrović, 2018) in the classical linear regression model is present in the case when the assumption of the constancy of the random error variance for all observations is violated, so in that case $Var(\varepsilon_i) = \sigma_i^2$, for each *i*. Applying least squares methods to a model with heteroskedastic errors leads to inefficient estimates. White's test is often used to detect heteroscedasticity in practice (White, 1980). Heteroskedasticity should be eliminated, if possible, by changing the specification or by including the influence of extreme observations using artificial variables, i.e. by transformation in order to obtain a constant variance, which leads to weighted least squares methods.

The Results

The analysis of the concentration of gross expenditures for R&D implies the inclusion in the analysis of the gross domestic product and the population of each observed country. It is a logical consequence of the interdependence between gross expenditure on R&D and the growth and development of a country (Komić, 2021).

Table 1 shows the values of the basic descriptive measures included variables for three years out of 25 observed.

Table 1

Descriptive Measures of the Included Variables

Descriptive measure	Year		
	1997	2010	2021
Population			
Arithmetic mean	70929976	79539276	85866981
Standard deviation	217313639	249246801	272927684
Variation coefficient	3.064	3.134	3.178
GDP			
Arithmetic mean	5.15471E+11	1.02825E+12	1.5494E+12
Standard deviation	1.35881E+12	2.36936E+12	4.04032E+12
Variation coefficient	2.636	2.304	2.608
GDP/cap			
Arithmetic mean	11941.5	23404.5	30521.5
Standard deviation	11486.7	19309.4	24022.5
Variation coefficient	0.962	0.825	0.787
GERD			
Arithmetic mean	10235.9	21930.3	40391.6
Standard deviation	34092.4	63283.8	127175.9
Variation coefficient	3.331	2.886	3.149
HDI			
Arithmetic mean	0.743	0.820	0.855
Standard deviation	0.094	0.079	0.077
Variation coefficient	0.127	0.096	0.090

Source: Edited by the author.

The variation intervals show extremely large differences in absolute values, which is logical to expect considering the different sizes of countries and differences in the level of development. It is also noticeable that the variability is extremely high for the absolute values of the variables, while this is not the case for the derived variables (GDP/cap and HDI). And this is a fact that points to the need to express the concentration using composite coefficients.

The data unequivocally indicate huge differences in terms of allocations for research and development, so this problem has not only a dynamic, but also a number of other dimensions. One of those dimensions is the concentration of research and development expenditures, which gives this analysis its full meaning. Considering the differences shown, it can be stated that the representativeness of the observed group of countries was significantly achieved in this way.

When looking at the average annual growth rates of the included variables, it can be seen that in the observed period from 1997 to 2021, there is a significant growth in gross expenditures for R&D and that it is more pronounced than in the case of gross domestic product, both absolutely and per capita. These are important positive trends, which indicate a growing appreciation of the importance of investment in research and development. These tendencies are significant both for individual countries and for humanity as a whole.

Table 2

Average Growth Rates of the Included Variables

Variable	Period		
	1997-2010	2010-2021	1997-2021
Population	0.885	0.698	0.799
GDP	5.455	3.798	4.692
GDP/cap	5.313	2.443	3.987
GERD	6.036	5.709	5.886
HDI	0.761	0.381	0.587

Source: Edited by the author.

A significant growth of the gross domestic product per capita can also be observed, which points to pronounced economic growth in the observed period. Among the population, growth rates are significantly lower, which points to the conclusion that changes in this dimension are nowhere near as pronounced as in terms of gross domestic product and gross allocations for R&D. And this is a fact that should be taken into account when deciding to include these two variables as criteria in the concentration analysis.

In addition to population size and gross domestic product, other dimensions of influence on concentration can be taken into consideration. These are variables that take into account other demographic segments, structural segments of the gross domestic product, as well as segments of the labor and business market, imports and exports, investments, macroeconomic results, financial and others. It is to be assumed that the size of the population, the gross domestic product and HDI sufficiently correctly express all the mentioned dimensions in a sublimated version, so in this work this was the approach.

Gini concentration coefficients can be calculated in several ways.

In this paper, three sets of concentration coefficients of gross expenditures for R&D in the observed time period were made.

Table 3 shows the values of three groups of concentration coefficients that were obtained: the first group without an additional criterion variable based on the number of inhabitants (Popul_noCV) and gross domestic product (GDP_noCV); the second group includes the criterion variable gross domestic product per capita (GDP/cap; Popul_GDP/cap; GDP_GDP/cap) and the third group of concentration coefficients is divided by including the criterion variable human development index (HDI; Popul_HDI; GDP_HDI). Each of the three previously indicated approaches has its advantages and disadvantages, and therefore the concentration coefficients were calculated in all three ways.

The first approach acknowledges the originality of the included variables that indisputably have a direct impact on gross expenditures on R&D. On the other hand, the question of comparability can be raised considering the large differences in the countries according to the number of inhabitants and the size of the gross domestic product. In the other two approaches, the effort goes in the direction of bringing the two involved variables into a logical connection and achieving comparability of the involved countries. In this sense, the variable gross domestic product per inhabitant, as already mentioned, has the power of a criterion variable, regardless of all its shortcomings. In a logical sequence, considering the properties of the previous criterion variable, the HDI is used as the criterion variable for the third way of calculating concentration coefficients.

Table 3

Concentration Coefficients According to the Size of the Population and GDP, According to the Level of GDP/cap and HDI

	Concentration coefficients					
	No criterium varuables		Criterium variable GDP/cap		Criterium variable HDI	
Year	Popul_noCV	GDP_noCV	Popul_GDP/cap	GDP_GDP/cap	Popul_HDI	GDP_HDI
1997	0.46715	0.16160	0.79409	0.17951	0.77839	0.14631
1998	0.46458	0.16560	0.79120	0.16866	0.78056	0.14394
1999	0.46340	0.15767	0.79788	0.18020	0.77972	0.13742
2000	0.45469	0.15694	0.79829	0.17502	0.77979	0.14201
2001	0.45344	0.15379	0.79061	0.15987	0.78067	0.14370
2002	0.45684	0.14489	0.78362	0.15047	0.77491	0.13468
2003	0.46659	0.14793	0.77456	0.14664	0.76858	0.13557
2004	0.47072	0.14649	0.76224	0.13250	0.75586	0.13167
2005	0.46632	0.14967	0.75397	0.13170	0.74798	0.13271
2006	0.46221	0.15348	0.74150	0.11781	0.74121	0.13672
2007	0.46200	0.15211	0.72317	0.09797	0.72963	0.14312
2008	0.45176	0.15969	0.71631	0.12490	0.72058	0.15847
2009	0.42561	0.14785	0.71227	0.14237	0.70592	0.14484
2010	0.41143	0.12817	0.69726	0.14365	0.68872	0.13845
2011	0.39880	0.12806	0.68064	0.13609	0.67622	0.14148
2012	0.37464	0.11775	0.66719	0.12292	0.66166	0.12856
2013	0.35682	0.12147	0.65364	0.11387	0.64925	0.12233
2014	0.34771	0.11471	0.64819	0.10947	0.63891	0.10471
2015	0.31797	0.11101	0.64700	0.10529	0.63876	0.11273
2016	0.31425	0.11102	0.64805	0.11055	0.63263	0.10146
2017	0.30891	0.11045	0.63997	0.11218	0.62606	0.10386
2018	0.30220	0.11274	0.64270	0.11460	0.61424	0.09315
2019	0.29886	0.12108	0.65350	0.13124	0.61992	0.10819
2020	0.28920	0.12427	0.66898	0.14448	0.61621	0.11336
2021	0.27713	0.12810	0.65678	0.14002	0.60659	0.11852
Growth rate	-2.152	-0.963	-0.788	-1.030	-1.034	-0.874

Source: Calculated by the author.

Already at first glance, differences in the obtained concentration coefficients can be seen, which are significantly higher in cases where they are calculated based on the population size. This is logical to expect, because the size of the gross domestic product in a peculiar way determines the amount allocated for research and development. However, the question that arises is whether, on the basis of the obtained concentration coefficients, unique concentration coefficients can be established. The answer is positive and points to the application of composite concentration coefficients.

Information on the relationships of the obtained concentration coefficients is given by the simple linear correlation coefficients whose values are given in Table 4 (lower triangular matrix).

In this analysis, three regression models are formed. In the first, the dependent variable is GERD, in the second, GDP/cap, and in the third, HDI. In all models, the independent variables are population size and gross domestic product. Based on the obtained models, three versions of the composite concentration coefficients are obtained by extrapolation, and their geometric mean gives the final version of the composite concentration coefficients.

Table 4

Coefficients of Simple Linear Correlation of Concentration Coefficients

	Popul_noCV	GDP_noCV	Popul_GDPcap	GDP_GDPcap	Popul_HDI	GDP_HDI
Popul_noCV	1.000					
GDP_noCV	0.884**	1.000				
Popul_GDPcap	0.891**	0.913**	1.000			
GDP_GDPcap	0.473*	0.640*	0.752**	1.000		
Popul_HDI	0.958**	0.906**	0.977**	0.636**	1.000	
GDP_HDI	0.850**	0.860**	0.743**	0.519**	0.791**	1.000

Notes: * ($p < 0.05$) ** ($p < 0.01$).

Composite coefficients of concentration in the subject analysis were obtained by extrapolation based on regression model scores:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i,\ i = 1, 2, ..., n$$

where is:

Y_i: *i*-th dependent variable (GERD in first model, GDP/cap in second and HDI in third)

$\beta_0, \beta_1, \beta_2$: model parameters

ε_i: stochastic term or random error

$k = 2$ is number of explanatory variables.

Independent variables are: X_1 – Population i X_2 – GDP.

The first information about the relationship between the included variables is provided by the simple linear correlation coefficients, the values of which are given in Table 5 (lower triangular matrix).

Table 5

Simple Linear Correlation Coefficients of the Included Variables

	Population	GDP	GERD	GDP/cap	HDI
Population	1.000				
GDP	0.919**	1.000			
GERD	0.977**	0.857**	1.000		
GDP/cap	0.924**	0.879**	0.820**	1.000	
HDI	0.992**	0.930**	0.952**	0.947**	1.000

Note: ** ($p < 0.01$).

All obtained simple linear correlation coefficients have high values and are highly statistically significant ($p < 0.01$). This is a fact that should be taken into account in further analysis.

The most important parameters of the estimated regression models are given in Table 6, Table 7 and Table 8.

Table 6

Parameters of the Regression Model and Analysis of the Presence of Multicollinearity (dep. var. GERD)

GERD	Non standardiz. coeff.		Standardiz. coeff.	*t*	*p*	Partial corr.	Tolerance	VIF
	Value	Std. error						
Constant	-7.8E+12	640.6E+9		-12.228	0.000			
Population	2271.247	185.179	1.224	12.265	0.000	0.934	0.155	6.454
GDP	-0.006	0.002	-0.269	-2.693	0.013	-0.498	0.155	6.454

Based on the obtained values, it can be concluded that there is no harmful influence of multicollinearity in the evaluated regression model. The coefficient of determination is very high and amounts to $R_i^2 = 0.966$, and its corrected value is 0.963. Analysis of variance ($F = 312.899$; $p = 0.000$) shows that the evaluated model as a whole is statistically significant, and from Table 6 it can be seen that the obtained coefficients of this model are also statistically significant. Based on the value of White's test ($\chi^2 = 25.000$; $p = 0.406$), it can be concluded ($p > 0.05$) that there is no heteroskedasticity in the model. The obtained results show that the obtained regression model is statistically significant as a whole and for individual parameters and that it can be used to obtain composite concentration coefficients.

Table 7

Parameters of the Regression Model and Analysis of the Presence of Multicollinearity (dep. var. GDP/cap)

GDP/cap	Non standardiz. coeff.		Standardiz. coeff.	*t*	*p*	Partial corr.	Tolerance	VIF
	Value	Std. error						
Constant	-5.9E+04	17580.993		-3.351	0.003			
Population	1.9E-05	0.000	0.746	3.673	0.001	0.617	0.155	6.454
GDP	5.5E-11	0.000	0.193	0.951	0.352	0.199	0.155	6.454

Based on the obtained values (Table 7), it can be concluded that there is no harmful influence of multicollinearity in the evaluated regression model. The coefficient of determination is high and amounts to $R_i^2 = 0.859$, and its corrected value is 0.846. Analysis of variance ($F = 67.175$; $p = 0.000$) shows that the evaluated model as a whole is statistically significant, while the coefficient with GDP is not statistically significant. Based on the value of White's test ($\chi 2 = 25.000$; p = 0.406), it can be concluded ($p > 0.05$) that there is no heteroskedasticity in the model. The obtained results show that the obtained regression model is statistically significant as a whole and that it can be used to obtain composite coefficients of concentration.

Table 8

Parameters of the Regression Model and Analysis of the Presence of Multicollinearity (dep. var. HDI)

HDI	Non standardiz. coeff.		Standardiz. coeff.	*t*	*p*	Partial corr.	Tolerance	VIF
	Value	Std. error						
Constant	0.247	0.034		7.190	0.000			
Population	1.3E-10	0.000	0.879	13.497	0.000	0.945	0.155	6.454
GDP	2.1E-16	0.000	0.122	1.875	0.074	0.371	0.155	6.454

Based on the obtained values, it can be concluded that there is no harmful influence of multicollinearity in the evaluated regression model. The coefficient of determination is very high and amounts to $R_i^2 = 0.986$, and its corrected value is 0.984. Analysis of variance ($F = 749.399$; $p = 0.000$) shows that the evaluated model as a whole is statistically significant, and from Table 8 it can be seen that the coefficient with GDP is not statistically significant. Based on the value of White's test ($\chi^2 = 25.000$; $p = 0.406$), it can be concluded ($p > 0.05$) that there is no heteroskedasticity in the model. The obtained results show that the obtained regression model is statistically significant as a whole and that it can be used to obtain composite coefficients of concentration.

The values of the obtained composite concentration coefficients are given in Table 9 and presented in Figure 1, together with gross expenditures for R&D per capita.

Integrated composite concentration coefficients as the final expression of all established relationships are obtained as the geometric mean of previously calculated composite concentration coefficients.

Discussion

Expressing the concentration with the Gini coefficient, in cases where there are several determinants, that is, possible criteria, results in obtaining different results. It is realistic to expect that some obtained results do not quantify the concentration in the right way.

Based on the results obtained and the analysis carried out, it can be concluded that the values of the obtained concentration coefficients are significantly different for different measurement criteria, which is very noticeable from Table 3. Based on the simple linear correlation coefficients between the obtained concentration coefficients, it can be concluded that all the obtained coefficients are statistically significant and only in two cases have relatively low values.

Table 9

Composite Concentration Coefficients Based on Regression Models

Year	Composite concentration coefficients			
	No criterium varuables	Criterium varuable GDP/cap	Criterium variable HDI	Integrated coefficients
1997	0.52846	0.62730	0.70229	0.61518
1998	0.52424	0.62305	0.70391	0.61262
1999	0.52493	0.63027	0.70238	0.61480
2000	0.51446	0.62957	0.70300	0.61064
2001	0.51378	0.62091	0.70398	0.60784
2002	0.52033	0.61388	0.69781	0.60632
2003	0.53145	0.60638	0.69236	0.60652
2004	0.53690	0.59445	0.68070	0.60116
2005	0.53065	0.58813	0.67389	0.59469
2006	0.52460	0.57614	0.66843	0.58677
2007	0.52471	0.55863	0.65903	0.57807
2008	0.51014	0.55871	0.65295	0.57093
2009	0.48131	0.55907	0.63839	0.55589
2010	0.46924	0.54811	0.62249	0.54300
2011	0.45380	0.53425	0.61187	0.52937
2012	0.42700	0.52167	0.59749	0.51056
2013	0.40418	0.50981	0.58581	0.49421
2014	0.39485	0.50489	0.57457	0.48565
2015	0.35943	0.50319	0.57542	0.47038
2016	0.35487	0.50499	0.56865	0.46709
2017	0.34849	0.49928	0.56317	0.46102
2018	0.33966	0.50178	0.55147	0.45467
2019	0.33333	0.51306	0.55830	0.45706
2020	0.32065	0.52717	0.55567	0.45457
2021	0.30484	0.51720	0.54784	0.44204
Growth rate	-2.266	-0.801	-1.030	-1.368

Source: Calculated by the author.

Calculating the composite concentration coefficients based on established regression models (Table 9) yields values where differences still exist, but are significantly smaller compared to the values of the concentration coefficients from Table 3. The difference in the composite concentration coefficients points to the need to integrate them to obtain a unique set of concentration coefficients which the author named integrated composite coefficient of concentration. This method enabled the sublimation of all concentration coefficients based on all included criteria.

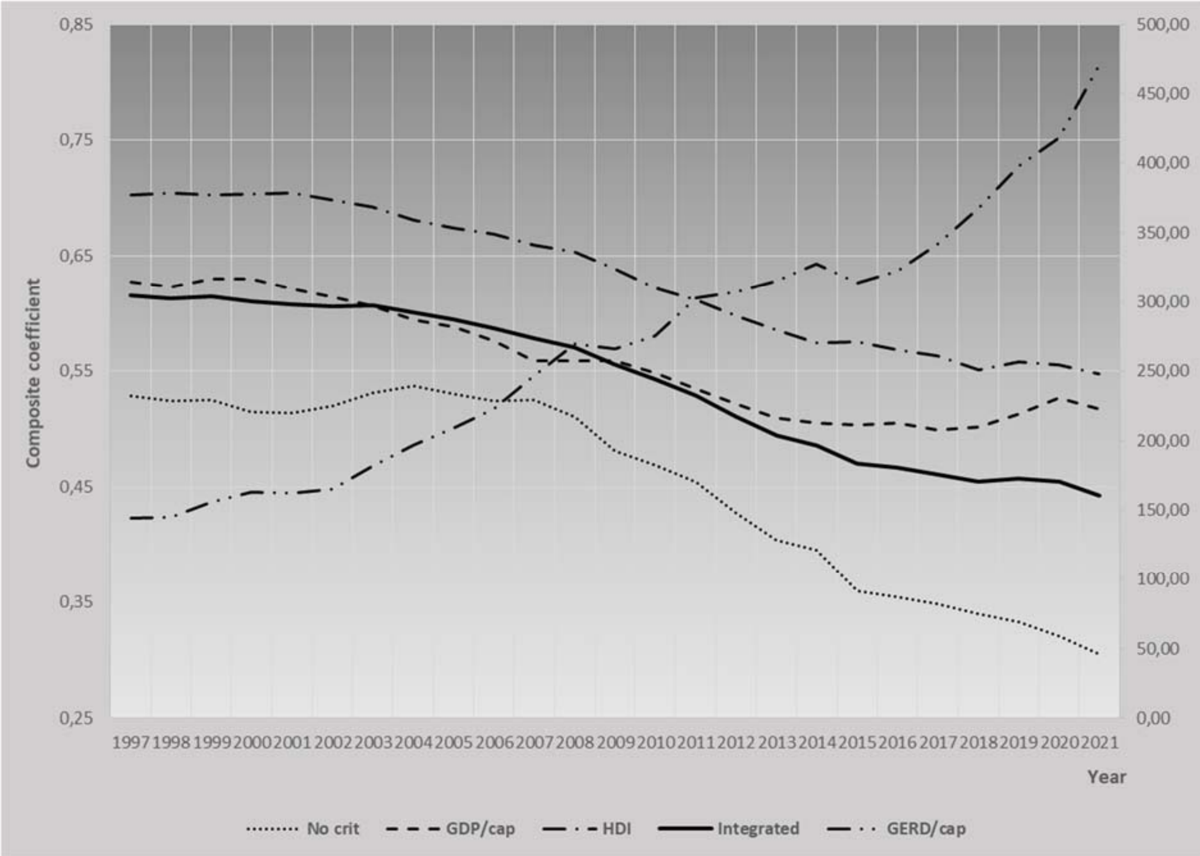

Figure 1. Composite concentration coefficients.

Table 10 shows the values of simple linear correlation coefficients (lower triangular matrix) between integrated and other composite concentration coefficients. From Table 10, it can be seen that the values of all obtained simple linear correlation coefficients are very high and that they are all statistically highly significant ($p < 0.01$). This clearly shows that the integrated composite concentration coefficients correctly sublimated the previously obtained composite concentration coefficients.

Table 10

Coefficients of Simple Linear Correlation of Composite Concentration Coefficients

	CC no Crit	CC_GDP/cap	CC_HDI	CC_Integrated
CC no Crit	1.000			
CC_GDP/cap	0.862**	1.000		
CC_HDI	0.957**	0.965**	1.000	
CC_Integrated	0.981**	0.942**	0.994**	1.000

Note: ** ($p < 0.01$).

Table 11 shows the values of the simple linear correlation coefficients of the integrated composite indices and the included variables expressed as relative values. There is a special emphasis on the variables that show the growth and development tendencies of the observed countries (GDP/cap and HDI).

Table 11

Coefficients of Simple Linear Correlation of Integrated Composite Concentration Coefficient and Included Variables

	CC_Integrated	GDP/cap	HDI	GERD/cap
CC_Integrated	1.000			
GDP/cap	-0.857**	1.000		
HDI	-0.942**	0.947**	1.000	
GERD/cap	-0.953**	0.926**	0.956**	1.000

Note: ** ($p < 0.01$).

Table 11 particularly clearly shows the meaning of this analysis: it is noticeable that the integrated composite coefficients of concentration are highly negatively correlated with all variables of a relative type that in a certain way express the development tendencies of the observed countries. This means that the development of the observed countries is accompanied by a decrease in the concentration of gross expenditure on R&D. Indirectly, this points to the fact that an increasing number of countries are investing more and more in research and development activities. The benefit for society as a whole cannot be left out.

Considering the dynamic aspect of the analysis, average growth rates provide additional information in this regard.

Based on the growth rates of the obtained concentration coefficients, it can be observed that all the obtained rates are negative. This leads to the conclusion that in the observed period, regardless of the method of calculating the concentration coefficients, it is evident that the concentration in terms of allocation for research and development is decreasing. Logically, it follows from this that the understanding of the importance of setting aside for research and increasing its absolute and relative aspects is increasing and is increasingly present in most countries in the world. It is to be expected that such tendencies will have a significant impetus for the growth and development of not only individual countries, but the human community as a whole.

In terms of gross expenditures for R&D in the global framework in the period from 1997 to 2021, there is a tendency of their absolute increase and increase of their relative participation in the gross domestic product, which points to positive tendencies in terms of development. These statements are confirmed by the data from Table 12. It is evident that the fastest growth in the observed period for the countries included in the analysis is precisely the gross expenditures for R&D per capita, while the growth of the share of gross expenditures for R&D in the gross domestic product is positive, but significantly less intense. At the same time, the growth rate of gross domestic product per capita is somewhere in the middle between the previous two rates. Another confirmation that the analysis of concentration with the inclusion of more variables and criteria has complete justification and sense.

A particularly important observation refers to two opposite dynamic tendencies: relative values of gross expenditures for research and development (share in gross domestic product GERD/GDP and in relation to the number of inhabitants GERD/cap), as well as gross domestic product per inhabitant tend to increase all concentration coefficients tend to decrease.

These two tendencies (partially presented in Chart 1) clearly point to the following conclusion: the evident growth of gross expenditures for research and development is accompanied by a decrease in their concentration in the set of observed countries in the observed time period.

Table 12

Relative Indicators of the Participation of Gross Expenditures for R&D in the Gross Domestic Product and in Relation to the Number of Inhabitants

Year	GERD/GDP%	GERD/cap	GDP/cap
1997	1.717	144.3	11941.5
1998	1.669	145.0	11897.0
1999	1.716	155.5	12142.8
2000	1.686	162.7	12065.4
2001	1.618	162.2	12050.1
2002	1.583	164.8	12795.5
2003	1.667	181.6	15031.8
2004	1.692	196.7	17210.3
2005	1.686	209.2	18629.1
2006	1.652	222.9	20229.9
2007	1.703	246.4	23248.8
2008	1.779	269.8	25653.8
2009	1.757	265.7	22630.6
2010	1.726	275.7	23404.5
2011	1.786	303.2	25842.7
2012	1.732	307.0	25229.8
2013	1.699	315.2	25965.8
2014	1.690	327.1	26197.5
2015	1.561	313.8	23263.0
2016	1.544	322.6	23377.2
2017	1.548	341.6	24958.8
2018	1.587	368.1	26865.3
2019	2.462	398.3	26630.5
2020	2.649	417.9	25743.5
2021	2.607	470.4	30521.5
Growth rate	1.755	5.047	3.987

Source: Calculated by the author.

Conclusion

Based on the conducted analyzes and obtained results, it can be concluded that the analysis of the dynamics of the distribution of gross expenditures for R&D in a relatively long period of time clearly quantifies the tendencies in this regard, which from the point of view of development flows is of particular importance. It can be concluded that there is an evident concentration of gross expenditures for R&D in the global context, which has significantly decreased in the observed period of time.

Expressing the concentration with the Gini coefficient can be performed with consideration of several possible criteria. In an analytical sense, this allows the present concentration to be viewed and measured from different angles. This leads to different results being obtained, as well as different (and wrong) conclusions regarding the concentration level. That is why it is necessary to take into account all dimensions of the problem, which primarily refers to the establishment of criteria and other included variables.

Establishing integrated composite concentration coefficients allows this problem to be solved. Composite concentration coefficients take into account all criterion variables and simultaneously express the concentration.

A particularly important observation refers to the fact that the relative values of gross expenditures for research and development (participation in the gross domestic product GERD/GDP and in relation to the number of inhabitants GERD/cap), as well as the gross domestic product per capita tend to increase, and that all coefficients concentrations tend to fall. This clearly points to the following conclusion: the evident growth of gross expenditures for research and development is accompanied by a decrease in their concentration in the set of observed countries in the observed period of time. Indirectly, this means that an increasing number of countries are increasing R&D expenditures, which means that the dispersion of research and development activities is increasing. This results in a far greater participation of all research and development capacities in solving existential, social, political, technological and all other problems of the human community.

In the observed period, positive tendencies in the dynamic view of gross expenditures for R&D were emphasized, both in absolute and relative terms. At the same time, there are evidently significant differences in the speed of dynamic changes in gross expenditures for research and development of the observed countries.

In several previous works (Komić, 2015, 2019), the author of this paper analyzed, among other things, the dynamic aspect of expenditures for research and development. In the conducted analytical procedures and conclusions, it can be stated that the established differences in the dynamics of gross allocations for R&D are extremely large. This means that humanity is deeply divided in terms of creating the conditions for development and that it is realistic to expect that the gap will continue to widen. However, aware that this is not for the benefit of the entire development flow of human society, the developed and highly developed must be aware of this fact, and in this sense, additional investments and joint efforts of both are needed for a kind of rapprochement that as a whole, and that means for everyone, they must and can give better results in the overall human development and all dimensions of life and work.

Monitoring and analysis of allocations for research and development presupposes the existence and availability of adequate, comprehensive, consistent and methodologically harmonized data, while respecting the dispersion and presence of inequality in the world community. In this regard, it is necessary to respect and constantly review the existing standards. as well as the introduction of new data collection, processing and presentation. In addition, it is necessary to define all the necessary indicators and indicators that need to be monitored and on the basis of which all analytical procedures can be adequately implemented and valid data can be obtained and consistent conclusions can be drawn.

References

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323-351. https://doi. org/10.2307/2951599

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis. *The Review of Economics and Statistics*, *49*(1), 92-107.

Field, A. (2018). Discovering statistics using SPSS (5rd ed.). Los Angeles: SAGE Publications.

Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal*, *31*(121), 124-126. https://doi:10.2307/2223319

Grossman, G. M., & Helpman, E. (1991). Trade, knowledge spillovers, and growth. *European Economic Review*, *35*, 517-526. https://doi. org/10.1016/0014-2921(91)90153-A

Hall, B. H., & Mairesse, J. (2009). Measuring the returns to R&D, national bureau of economic research. Working Paper No. 15622.

Janjić, V. (2013). Obrazovanje, nauka i proizvodnja hrane, Akademija nauka i umjetnosti Republike Srpske, Banja Luka.

Jones, B. F., & Summers, L. H. (2020). A Calculation of the Social Returns to Innovation, National Bureau of Economic Research, https://doi 10.3386/w27863

Jovičić, M. (2002). Ekonometrijski metodi, Centar za izdavačku djelatnost, Ekonomski fakultet Beograd.

Jovičić, M., & Dragutinović-Mitrović, R. (2018). Ekonometrijski metodi i modeli, Centar za izdavačku djelatnost, Ekonomski fakultet Beograd.

Komić, J. (2021). Ulaganje u istraživanje i razvoj—investicija za budućnost, Univerzitet u Banja Luci, Ekonomski fakultet, Banja Luka.

Komić, J. (2015). Ulaganje u istraživanje i razvoj—investicija za budućnost, Akademija nauka i umjetnosti Republika Srpske, Naučni skupovi, Uloga akademija nauka i umjetnosti u razvoju društva budućnosti, Banja Luka, Zbornik radova, pp.111-130.

Komić, J. (2019). Istraživanje i razvoj u malim zemljama u uslovima globalizacije, Razvoj malih zemalja u uslovima globalizacije, Akademija nauka i umjetnosti Republike Srpske, Naučni skupovi, Knjiga XLVII, pp. 529-552.

Komić, J. (2019) Some aspects of quantitative analysis of development role of gross expenditures for research and development (2019). *ICOS—International Conference on Official Statistics 2019*, Book of Abstracts, Sarajevo, pp. 24-25.

Kovačić, B., Opačić, R., Marohnić, L., O Ginijevu koeficijentu koncentracije, math.e Hrvatski matematički elektronički časopis, p.10 (Retrieved, March, 2021).

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, *9*(70), 209-219. https://*doi:10.2307/2276207*

Lovrić M., Komić J., & Stević S. (2017). Statistička analiza—metodi i primjena, II izdanje, Narodna i univerzitetska biblioteka Republike Srpske, Banja Luka.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3-42. https://doi.org/10.1016/0304-3932(88)90168-7

Menard, S. (2004). Standardized regression coefficients, in Lewis-Beck, M. S., Bryman, A., Liao, T. F. (eds.), The Sage Encyclopedia of Social Science Research Methods, Sage Publishing, pp. 1069-1070. https://doi:10.4135/9781412950589.n959

Nekrep, A., Strašek, S., & Boršič, D. (2018). Productivity and economic growth in the European Union: Impact of investment in research and development. *Naše gospodarstvo/Our Economy*, *64*(1), 18-27. https://doi: 10.2478/ngoe-2018-0003

OECD (2002). Frascati Manual 2002: Proposed standard practice for surveys on research and experimental development. *The Measurement of Scientific and Technological Activities*, OECD Publishing, Paris.

OECD (2015). Frascati Manual 2015: Guidelines for collecting and reporting data on research and experimental development. *The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264239012-en

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5), 71-102. https://doi.org/10.1086/261725

Stanton, E. (2007). The human development index: A history, political economy research institute. Working Papers no. 127, University of Massachuttes Amherst. https://doi.org/10.7275/1282621.

Somun-Kapetanović, R. (2008). Statistika u ekonomiji i menadžmentu, II izdanje, Ekonomski fakultet u Sarajevu, Sarajevo.

Šošić, I. (2004). Primijenjena statistika, 2. izdanje, Školska knjiga, Zagreb.

UNDP (1990). *Human Development Report 1990.* Oxford University Press. New York. https://hdr.undp.org/system/files/documents/hdr1990encompletenostatspdf.pdf

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838.