RELATIONSHIP BETWEEN EDUCATION AND ECONOMIC GROWTH: A COMPARATIVE STUDY BETWEEN ROMANIA AND POLAND

ANDREEA OANA BEGU
BOGDAN ZAINEA
MARIA-CRISTINA ZORILĂ
The Bucharest University of Economic Studies
ROMANIA

Abstract: Through this paper, we aim to address a topic whose origins date back to the 1960s, when the economist Theodore W. Schultz laid the foundations of human capital theory and argued that: if both physical capital and human capital develop together, then economic capital can also grow (T.R. Breton, 2014). In line with this statement, the present study aims to highlight the extent to which certain indicators, closely related to education, are connected to a country's economic growth. Our analysis will focus on two Eastern European countries that have undergone a period of both economic and social transition, namely Romania and Poland. The main hypothesis from which we start our research is that the educational component of human capital has a positive relationship with economic growth. Thus, in the first part of the paper, we will present a comparative overview of the values for the selected indicators of education and economic growth for both countries. Subsequently, we will test two more auxiliary hypotheses, based on the primary one, through simple linear regression models, which will provide some interesting perspectives.

Keywords: education, economic growth, simple linear regression models, Romania, Poland

1. Introduction

Throughout history, education has been perceived and used in different ways, reflecting the social structures and economic needs of each era. Since ancient times, in societies such as Egypt and Mesopotamia, education was seen as a privilege reserved for the upper classes, managed by priests or ruling authorities. In this context, learning meant not only acquiring basic skills such as writing and reading, but also access to advanced knowledge such as law, medicine, religion or astrology, according to the standards of the time. The educational process often took place in knowledge centers, such as libraries, and pedagogical methods were based on memorization, rewriting or oral repetition.

Moving on to the classical periods, such as Ancient Greece and the Romans, we see a gradual opening up of education to the middle classes and not just to the elite. At the same time, the focus begins to shift from religion and manuscripts to oratory and the training of jurists, as in ancient Rome, or to critical thinking, philosophy and art in Greece. An interesting aspect of this period is the emphasis on physical education, especially in Sparta, where the training of citizens was seen almost exclusively from a military perspective.

While these historical developments laid the foundation for how education is structured and valued, it is important to note that its form and function have continued to evolve up, often in different directions, depending on the country, cultural background, political system, or religious influence. In today's world, education is no longer a uniform process, but one that adapts to the needs and aspirations of both individuals and societies.

In this paper, we approach education from our own contemporary perspective, which sees it as the process through which knowledge and skills are transmitted by qualified professionals, in formal or informal settings, including public, private, or online environments. This process helps individuals gain clarity on how to shape their personal and professional paths, aligned with their needs and ambitions. Thus, through the process of education, both formal and self-directed, a country prepares competent individuals who are ready to be integrated into the labor market, contributing to the creation of a foundation for sustainable and balanced economic growth.

From this definition and from the historical examples illustrated above, we observe a common aspect which is that: education has always been shaped by, and in turn helps shape, the socio-economic context in which it operates. This leads us to the central question of our paper: what is the relationship between education and the economy, and how do they influence each other in contemporary society?

2. Literature review

In this context, starting from the early ’60s, interest in a possible connection between education and economic growth began to bloom significantly. Among the pioneers of this field of study, we can highlight the economist Theodore W. Schultz, one of the founders of human capital theory, who argued the following: the economic growth of a country can occurs only if physical capital and human capital grow together. (T. R Breton, 2014; T.W. Schultz,1961).

To better understand what Schultz intended to say in the cited statement above, we first need to familiarize ourselves with the terms he used in this 'definition'. In this regard, Romanian sociologist Bogdan Voicu, in his comprehensive analysis of human capital, offers an extended definition compared to Schultz’s, one that captures the true essence of the concept “human capital”. He emphasizes that it is composed of the totality of personal skills each individual possesses, which retain their utility regardless of the social context in which they are placed (B. Voicu, 2004). Regarding the second part of Schultz’s definition, more precisely the economic growth, it refers to a country’s capacity to stimulate the production of goods and services over a specific period of time. Traditionally, this growth is measured using a classic aggregate indicator: Gross Domestic Product (GDP), in its various forms.

Now that both concepts from Theodore W. Schultz’s statement are clearly defined, we took a closer look at them and explored the connection between the two.

Starting with human capital and returning to Bogdan Voicu’s study, the two main components of human capital are educational capital and biological capital. The first is represented by the educational process described earlier, while the second refers to an individual's health status (B. Voicu, 2004; T. W. Schultz, 1961). In line with the main theme of this study, we have chosen to focus solely on the educational component, leaving the opportunity for future research to explore the second component as well, in parallel with economic growth, as previously mentioned, represented by GDP and its variations.

This niche area of study, concerning the relationship between education, previously referred to as educational capital, and economic growth, has been intensively researched over the past six decades, highlighting new working hypotheses and diverse conclusions. However, for the purpose of our analysis, we have chosen to focus on those studies that examine investments in education, under various forms, and other relevant educational indicators in relation to economic growth. These studies served as the foundation for the methodology and hypotheses formulated in our analysis. But before addressing these aspects, we invite you to review several of the key studies identified below.

A significant starting point in this field of research is represented by Barro’s studies from 2002 and 2013, as well as those conducted in collaboration with Lee (Barro & Lee, 1993; Barro & Lee, 1996), in which the considerable influence of education on economic growth was emphasized by comparing approximately 100 countries over a specific time period. Returning to Barro’s individual work, he demonstrated that there is a direct causal relationship from education, measured through school enrollment rates, and that education has a positive effect on economic growth (Barro, 2002; Barro, 2013; Bogdan Oancea et al., 2017)

These studies intrigued many other researchers to analyze more deeply the causal relationship between education and economic growth, focusing on their own countries. A good example from Europe, more precisely from Greece, is represented by Agiomirgianakis et al. (2002), who measured education through enrollment rates in primary, secondary, and tertiary education, in relation to economic growth. Their results indicated a direct causal relationship for the first two levels of education, and a reverse causal relationship for the third (Bogdan Oancea et al., 2017). In Asia, there are also many similar studies. A few examples include: the study by Huang et al. (2009), which found that in China, between 1972 and 2007, there was a positive relationship between the tertiary education enrollment rate and economic growth; and the study conducted by Hussin and Muhammad (2012), which emphasized that educational capital positively influences economic growth in Malaysia (Ștefan Cristian Ciucu, 2015).

Transitioning toward the studies that serve as direct foundation and inspiration for our methodology. We begin with the work conducted by Bogdan Oancea et al. (2017). In their research, they carried out a comparative analysis between Romania and the Czech Republic, two countries that experienced a transitional period, both economically and socially. The study illustrates the causality and long-term relationship between education, measured by the number of students enrolled in tertiary education per 100,000 inhabitants, and government expenditure on tertiary education, and economic growth. Their findings indicate a positive influence of the selected education-related indicators on economic growth during the analyzed period.

The second study, which, alongside the previous one, helped us define our working hypothesis, is authored by Ștefan Cristian Ciucu and Raluca Dragoescu (2015). They conducted a parallel analysis of Romania and Hungary, again focusing on countries that went through a transitional period, on the same topic. Education was represented by the number of persons enrolled in primary, secondary, and tertiary education, as well as by government investment in the education sector, measured as percentage of GDP, in comparison with real GDP, which reflects economic growth. This study found both short-term and long-term relationships between the mentioned variables, emphasizing that the higher number of individuals enrolled in tertiary education is a positive indicator of an evolving economy.

3. Methodology

Based on the studies discussed above, particularly the last two, the following section outlines the research methodology, the variables employed, the main hypotheses, and the expected results. This comparative study of Romania and Poland over the period 2000–2022 investigates whether education directly influences the economic growth. We expect this hypothesis to be positive.

For this purpose, education is operationalized as general government expenditure on education as a percentage of GDP, whereas economic growth is measured by the annual GDP growth rate. In addition to these two metrics, tertiary school enrollment rate, as percentage, was employed in the stationarity and multicollinearity tests.

The origin of the data for each variable is presented below:

General government expenditure on education, expressed as a percentage of GDP – Eurostat.

Annual GDP growth rate – Eurostat.

Tertiary school enrollment rate – World Bank.

Regarding the methodology applied to these metrics, the primary objective is to conduct a time-series simple regression between general government expenditure on education and annual GDP growth. To ensure the validity of the results, a series of preliminary tests were performed prior to the regression. The statistical tests employed are listed below, together with their general purpose:

Heteroscedasticity Test:

This kind of test examines whether the variability of the errors differs across values of the explanatory variable, or whether the error variance is constant across observations (homoscedasticity).

The specific test employed in this study is the Breusch–Pagan test, which assesses whether the variance of the residuals depends on the explanatory variables. This test involves running an auxiliary regression that includes the residuals obtained from the main regression and the explanatory variable. Based on this regression, we can calculate the LM (Lagrange Multiplier) statistic as $LM = n \times R^2$, where n is the number of observations and R^2 is the R square from the auxiliary regression. If $LM > \chi^2(\alpha, df)$ (the critical value), where α is the significance level and df the degrees of freedom, then we accept the alternative hypothesis H1: the errors are heteroscedastic, else we accept the null hypothesis H0: the errors are homoscedastic.

Thus, the heteroscedasticity test verifies whether the regression errors have constant variance, that is, whether homoscedasticity is present. If not, the standard errors and p-values are no longer reliable, and in this case, corrections need to be applied.

Autocorrelation Test:

This kind of test verifies whether the errors are independent from each other. If autocorrelation is present, the coefficients become inefficient and the standard errors are no longer valid, which means there is a risk of drawing incorrect conclusions about the significance of the coefficients. To test for this, the Durbin–Watson (DW) test and the Breusch–Godfrey test are commonly used (the latter being applied when testing for autocorrelation in an ARDL model).

If a DW test is applied to a simple model without lags and it indicates the presence of autocorrelation ($DW > 2$ or $DW < 2$), the issue can be corrected by estimating an ARDL model, where one or more lags of the variables are included (introducing more lags reduces the number of observations, which in turn increases the p-value).

On this model, the Breusch–Godfrey test is then applied. Similar to the Breusch–Pagan test, the LM statistic is calculated, from which the p-value = f(LM, df) is obtained, where df = the number of lagged residuals. If p-value > 0.05, the null hypothesis H0 (no autocorrelation) is accepted; otherwise, the alternative hypothesis H1 (autocorrelation is present) is accepted.

Stationarity Test:

This kind of test verifies whether the statistical properties of a series do not depend on time (constant mean and variance, and autocovariance depending only on the distance between observations rather than on the specific point in time).

Thus, if the values of a series fluctuate around a constant mean and the variations have the same amplitude, the series is stationary. However, if the series increases steadily over time, it is non-stationary (it has a trend). When a series is non-stationary, it can lead to misleading results (spurious regressions). Therefore, stationarity is important to ensure the stability of the estimated relationships and to demonstrate that they do not depend on the particular years chosen. In this regard, it was verified whether the mean, variance, and autocovariance of the series remain constant over time.

Multicollinearity Test:

Multicollinearity refers to the situation where two or more explanatory (independent) variables are highly correlated with each other (ex: $r > 0.8$). In short, this test checks whether the independent variables in the model are too strongly correlated, which would make the estimated coefficients unstable, difficult to interpret, and associated with inflated standard errors. As a result, variables may appear statistically insignificant even though they seem important.

In this case, the aim was to analyze the tertiary education enrollment rate to verify whether it can be introduced as an independent variable alongside government expenditure on education, and to check whether these two variables exhibit a high correlation. To perform this test, the correlation matrix between the two explanatory variables is constructed. If the correlation value exceeds 0.8, the variables are considered correlated and cannot be used simultaneously in the regression. If the correlation is below 0.8 but still relatively high, raising suspicions of multicollinearity, the Variance Inflation Factor (VIF) indicator is used. If $VIF < 5$, multicollinearity is not a concern; if $VIF > 10$, there is severe multicollinearity, and the coefficients become unreliable.

Based on these tests, it is possible to determine whether the DLM (Distributed Lag Model) or the ARDL (Autoregressive Distributed Lag Model) is more appropriate for this study.

The DLM is a regression model that captures the lagged effects of an explanatory variable on the dependent variable. Its equation is:

$$Yt = \alpha + \beta 0Xt + \beta 1Xt - 1 + \beta 2Xt - 2 + \cdots + \beta kXt - k + ut$$

Y_t = the dependent variable at time t (in this study, represented by the annual GDP growth rate, expressed as a percentage);
$X_t, X_{t-1}, \ldots, X_{t-k}$ = the current and lagged values of the independent variable (in this case, government expenditure on education);
$\beta_0, \beta_1, \ldots, \beta_k$ = the lag coefficients, i.e., the extent to which the current and past values of the explanatory variable influence the dependent variable;
u_t = the error term;
Long-run effect = $\beta_0 + \beta_1 + \ldots + \beta_k$ (the total long-term effect of the explanatory variable).

The ARDL is a more general version of the DLM which, in addition to the lagged effects of the explanatory variable (government expenditure on education as a % of GDP), also includes the lagged effects of the dependent variable (annual GDP growth rate). An advantage of this model is that it captures the fact that the current value of the dependent variable depends not only on the explanatory variable but also on its own past values. The ARDL is often used when the aim is to study both short-run and long-run relationships between variables (especially in economic time series). Its equation is:

$Yt=\alpha+\phi 1Yt-1+\cdots+\phi pYt-p+\beta 0Xt+\beta 1Xt-1+\cdots+\beta qXt-q+ut$

Y_t = the dependent variable at time t (in this study, represented by the annual GDP growth rate, expressed as a percentage);
$X_t, X_{t-1}, \ldots, X_{t-k}$ = the current and lagged values of the independent variable (in this case, government expenditure on education);
$\beta_0, \beta_1, \ldots, \beta\_q$ = the lag coefficients of the explanatory variable, i.e., the extent to which the current and past values of the explanatory variable influence the dependent variable;
$\phi_1, \ldots, \phi\_p$ = the lag coefficients of the dependent variable, i.e., the extent to which the current value of the dependent variable is influenced by its own past values;
u_t = the error term;

Long-run effect= $\frac{\beta 0+\beta 1+\cdots+\beta q}{1-(\phi 1,+\ \ldots+\ \phi p\ )}$, which captures the long-term impact of the explanatory variable, taking into account that the dependent variable also depends on its own past values, meaning that the effects of the explanatory variable propagate both directly and through the dynamics of the dependent variable;

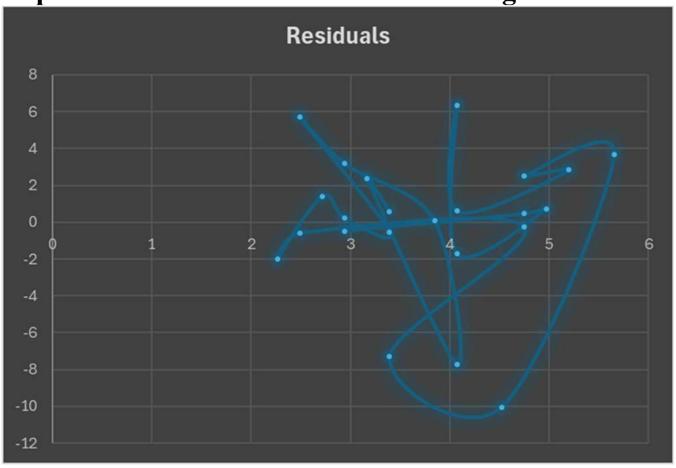
4. Results

This section presents the results obtained from the methodology described above, applied to the selected data, together with their interpretation and relevance to the main working hypothesis.

Starting off with the Breusch-Pagan test, which was used for evaluating the presence of heteroscedasticity in the desired regression model. The two hyphotesis for this test were the following:

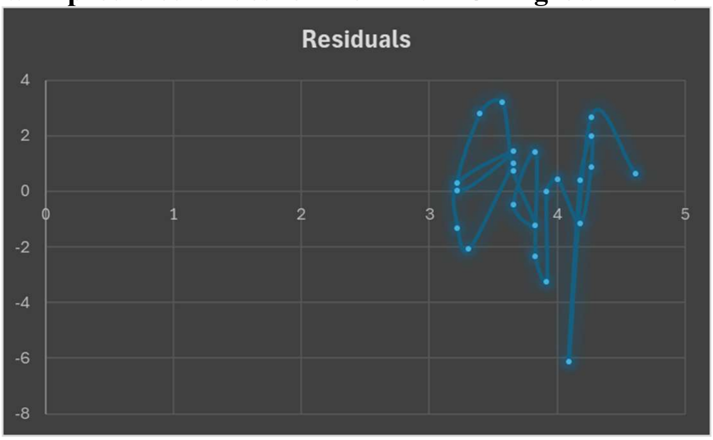
H0: it exists omoscedasticity.
H1: it exists heteroscedasticity.

Figure 3.1 –The graph with predicted values for the annual GDP growth rate and residuals for Romania



In the case of Romania, the figure 3.1, which contain the graph plotting the predicted values of annual GDP growth against the residuals from the regression of government expenditure on education on GDP growth suggests a slight fan-shaped heteroscedasticity. Nevertheless, the Breusch–Pagan test indicates that LM = 0.336812, which is below the critical value of $\chi^2(0.05;1) = 3.841459$. Thus, the null hypothesis cannot be rejected. The test was performed at a 5% significance level, equivalent to a 95% confidence level, ensuring the robustness of the results.

Figure 3.2 –The graph with predicted values for the annual GDP growth rate and residuals for Poland



In the case of Poland, the graph from figure 3.2 displays similar patterns. Here too, the LM statistic is lower than the critical value, with a value of 0.32802. Consequently, the null hypothesis is not rejected, which suggests that the dispersion of errors remains constant across all observations.

The following test applied to analyze the dynamics of the two variables under this study is the autocorrelation test, conducted using the Durbin–Watson (DW) statistic. The formula is given by: $DW=\frac{\sum(Ut-Ut\quad)^2}{\sum Ut^2}$, where:
Ut = regression's residuals.
Ut-1 = the lag coefficient of residuals.

For Romania, the calculation above yields a DW value of 1.740334, which is less than 2, indicating positive autocorrelation. To address this issue, a new regression is estimated using an ARDL model that incorporates one lag for the dependent variable and one for the explanatory variable. In this context, the Breusch–Godfrey test is applied, since the DW statistic cannot be used with ARDL models that include lagged dependent variables.

The hypotheses are defined as follows:
H0: autocorrelation doesn't exist.
H1: autocorrelation exists.

To test these hypotheses, an initial regression is estimated using the ARDL model, from which a new set of residuals is obtained. These residuals are then used in an auxiliary regression including their lags, along with the lags of both the dependent and explanatory variables, as well as the explanatory variables themselves.

This procedure yields R^2 =0.0024, which is then used to compute LM=0.0505, and a corresponding p-value of 0.824779. Since the p-value is greater than 0.05, the null hypothesis cannot be rejected. This indicates that no autocorrelation is present, and the standard errors are therefore valid.

In the case of Poland, no issues were encountered in applying the DW test. The test produced a value of 2.01979, differing from 2 by only 0.01979, which indicates no evidence of autocorrelation. Consequently, the residuals are not correlated with those of previous years and do not bias the regression model.

After obtaining the positive results for both countries following the application of the heteroscedasticity and autocorrelation tests, which were carried out using the basic model (annual GDP growth rate and government expenditure on education), since these tests verify the assumptions of the estimated specification, the next tests are those of multicollinearity and stationarity. However, the variable "tertiary school enrollment rate" was additionally introduced to test for potential multicollinearity with government expenditure on education, given that the two may be strongly correlated.

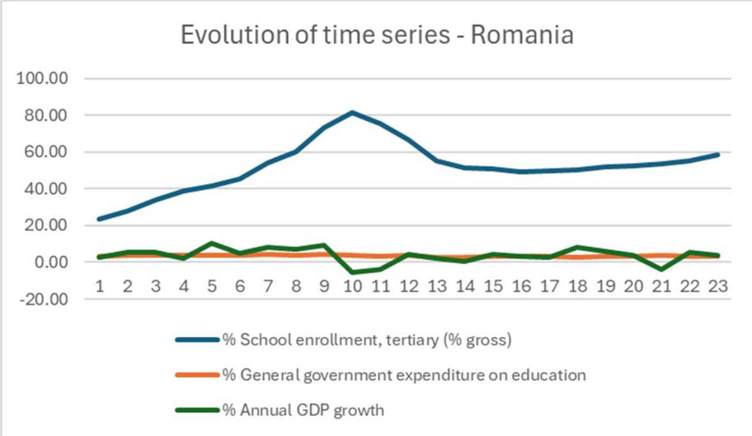
The results indicate a weak positive correlation in the case of Romania, with a coefficient of 0.16225655. This suggests that the variables move in the same direction but do not strongly influence each other, meaning they can be used simultaneously as independent variables in the regression model without multicollinearity issues. In the case of Poland, minor complications arose, as the correlation coefficient was –0.621290475, relatively close to the critical threshold of –0.80. To verify the potential multicollinearity issue, the Variance Inflation Factor (VIF) was calculated using the formula $VIF=\frac{1}{1-R^2}$.

The results yielded values of approximately 1.62 for both variables (government expenditure on education and tertiary school enrollment rate). Since these values are well below the critical threshold of 5, the results indicate the absence of multicollinearity.

Finally, a stationarity test was applied to the time series in order to verify that the data are not time-dependent and to avoid spurious results. Since no specialized statistical software was available to perform the Augmented Dickey–Fuller (ADF) test, alternative methods were employed. Specifically, the dataset was divided into two subperiods (2000–2010 and 2011–2022), and the mean and variance of each were calculated and compared. If the values were similar across the two subperiods, the series was considered stationary; if they differed substantially, the series was considered non-stationary.

For Romania, government expenditure on education and the annual GDP growth rate displayed similar means and variances across both subperiods, suggesting that these series are stationary. By contrast, the tertiary school enrollment rate showed significant differences in variance, indicating non-stationarity. In addition, correlation was computed within an ARDL model including the explanatory variables (tertiary enrollment and government expenditure on education), the dependent variable (GDP growth), and the lags of both expenditure and GDP growth. Since most coefficients were below 0.8, the results confirmed no risk of non-stationarity.

Figure 3.3 – The evolution of the 3 selected variables in Romania

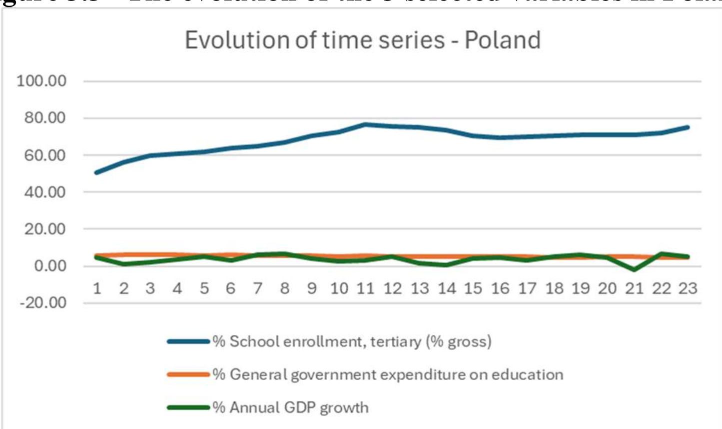


The graph from figure 3.3 was also produced to visualize the trends. It shows that both the GDP growth rate and government expenditure on education fluctuate around a stable mean, while the tertiary enrollment rate exhibits large variations with differing amplitudes. Considering the evidence from these three approaches, we can conclude that government expenditure on education and GDP growth are stationary, whereas the tertiary enrollment rate is not. To avoid spurious regressions, the tertiary enrollment variable was excluded from the final analysis, despite the absence of multicollinearity with government expenditure on education.

For Poland, similar results were obtained, with minor differences. In the comparison of means and variances across the two subperiods, a significant difference was again observed for the tertiary school enrollment rate, while a smaller difference appeared in the variance of the annual GDP growth rate. However, this difference was not large enough to render the GDP growth variable non-stationary.

Regarding the ARDL model, higher correlations were observed between tertiary school enrollment and government expenditure on education, as well as between the lag of government expenditure and expenditure itself. Nevertheless, since only a few coefficients displayed high values, the correlation matrix alone could not determine whether the series were non-stationary.

Figure 3.3 –The evolution of the 3 selected variables in Poland



A final element that helps clarify the results is the graphical analysis of the series, represented in the figure 3.4. As in the case of Romania, the tertiary school enrollment rate shows large fluctuations with varying amplitudes, while the other two variables evolve relatively steadily around a mean. Consequently, the tertiary enrollment variable was again excluded from the analysis in order to avoid biasing the regression results.

5. Conclusions

The economic analysis examined whether government spending on education directly influences annual GDP growth, being a comparison between Romania and Poland. The methodology included a number of tests, such as heteroscedasticity, autocorrelation, multicollinearity and stationarity, to ensure the validity of the results.

Their results indicated that both data sets, government spending on education and annual GDP growth rate, are adequate for analysis, with only the variable 'tertiary education enrolment rate' being eliminated, as it has been shown to be unstationary and may lead to erroneous results. In terms of autocorrelation, the results differed between the two countries, the problems identified in the case of Romania being solved by using an ARDL model. In contrast, the heteroscedasticity and multicollinearity tests showed for both countries that there were no major problems, allowing the simultaneous use of variables. Thus, the tests carried out justify the use of DLM and ARDL models to analyze the relationship between government spending on education and the annual GDP growth rate in the long and short term.

Further discussion: As an extension of the present analysis, several directions for future research may be pursued in order to uncover hidden patterns among relevant sets of indicators. Specifically, the comparative methodological framework employed here, which involved testing the adequacy of the Distributed Lag Model (DLM) versus the Autoregressive Distributed Lag (ARDL) model and interpreting the results, could be applied to alternative indicator pairs that may further illuminate the relationship between education and economic growth in Romania and Poland. A few possible avenues include:

- GDP per capita as a measure of economic growth, in relation to general government expenditure on education (as a percentage of GDP), which serves as a proxy for public investment in human capital.
- GDP per capita in relation to population by sex, age, and educational attainment (ages 15–64), with a focus on individuals who have completed higher education.
- GDP per capita in relation to number of researchers in R&D per one million inhabitants, which not only reflects advanced educational attainment but also indicates an economy's capacity for innovation, productivity enhancement, and international competitiveness.

The inclusion of the latter educational indicator, namely the number of researchers per one million inhabitants, is particularly noteworthy, as it provides an additional explanatory dimension for economic growth. By capturing both the qualitative outcomes of higher education and the innovation potential of the labor force, this variable may yield deeper insights into the mechanisms through which education contributes to long-term economic dynamics.

References:

[1] Theodore R. Breton (2014), "A human Capital theory of Growth: New Evidence for an Old Idea".

[2] Theodore W. Schultz (1961), "Investment in human Capital", The American Economic Review, Vol 51, No. 1 (Mar.1961), pg 1-17.

[3] Bogdan Voicu(2004)," Capitalul Uman: Componente, Niveluri, Structuri. Romania in Context European".

[4] Barro, R. J. (2002). Education as a Determinant of Economic Growth, in Lazear, E. P., eds., Education in the Twenty-first Century, Palo Alto, The Hoover Institution, 9–24.

[5] Barro, R. J., Lee, J. W. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. Journal of Development Economics, 104(C), 84–198.

[6] Bogdan Oancea, Richard Pospíšil, Raluca Mariana Drăgoescu (2017), "Higher Education and Economic Growth. A Comparison between Czech Republic and Romania".

[7] Agiomirgianakis, G., Asteriou, D., Monastiriotis, V. (2002). Human Capital and Economic Growth Revisited: A Dynamic Panel Data Study. International Advances in Economic Research, 8(3),177–187.

[8] Stefan Cristian Ciucu, Raluca Dragoescu,2015," The Relationship Between Economic Growth and Education: The Case of Romania and Hungary".

[9] Huang, F., Jin, L., Sun, X. (2009). Relationship between Scale of Higher Education and Economic Growth in China. Asian Social Science, 5(11), 55–60

[10]Hussin, M.Y., Muhammad, F. (2012), "Education Expenditure and Economic Growth: a Causal Analysis for Malaysia", Journal of Economics and Sustainable Development, Vol. 3, No. 7.

[11]Barro & Lee (1993) – "International Comparisons of Educational Attainment"

[12]Barro & Lee (1996) – "International Measures of Schooling Years and Schooling Quality"