

INTERDEPENDENCIES BETWEEN EDUCATIONAL ACHIEVEMENT, INCOME, AND UNEMPLOYMENT IN ROMANIAN COUNTIES

Emanuelle PERȚA*, Monica ROMAN**, Iliana CARAGEA**

*“Nicolae Bălcescu” Land Forces Academy, Sibiu, Romania

**Bucharest University of Economic Studies, Romania

perta.emanuelle@armyacademy.ro

Abstract: *This study examines the relationship between Baccalaureate pass rates, unemployment rates, and average income levels across Romanian counties in 2023. Using secondary statistical data from official national sources, we applied a range of quantitative methods, including bivariate correlations, multiple linear regression, and spatial econometric modeling, to explore these socio-economic interdependencies. Findings indicate a moderate but statistically non-significant negative correlation between pass rates and unemployment, suggesting that higher educational achievement may be associated with lower unemployment, though this relationship lacks robust statistical support. In contrast, average income levels emerged as a significant negative predictor of unemployment in the multiple regression model. Spatial regression analysis further revealed substantial spatial autocorrelation in unemployment rates, indicating that neighboring counties tend to exhibit similar labor market dynamics. These results highlight the importance of integrating both socio-economic and spatial dimensions into policy frameworks aimed at reducing regional disparities in employment and fostering equitable development.*

Keywords: educational achievement, grades, unemployment, graduation rate, socioeconomic status

1. Introduction

In recent years, the intricate relationship between educational attainment and labor market outcomes has garnered increased attention within the realm of social and economic research. This study embarks on a comprehensive exploration of the causality between the Baccalaureate graduation rate and the unemployment rate across various counties in Romania. The Baccalaureate, a pivotal milestone in the Romanian education system, serves as a gateway to higher education and potentially influences an individual's employability. Understanding the dynamics between Baccalaureate success rates and unemployment is crucial for policymakers,

educators, and economists seeking to formulate effective strategies for human capital development and labor market integration.

The main objective of this study is to investigate the causal relationship between the Baccalaureate graduation rate and the unemployment rate at the county level in Romania. By analyzing the statistical and econometric patterns, the research aims to provide insights into how variations in educational achievement may impact the prevalence of unemployment, shedding light on potential policy implications and interventions.

The next two research hypotheses were formulated:

Causality from Education to Employment: The study hypothesizes that a higher Baccalaureate graduation rate will lead to a subsequent reduction in the unemployment rate. This assumes that a well-educated workforce is better equipped to meet the demands of the labor market, thereby lowering unemployment.

Reverse Causality: Alternatively, the study will explore the possibility of reverse causality, where a higher unemployment rate might influence Baccalaureate graduation. Economic conditions, such as a challenging job market, could impact educational outcomes, affecting student motivation and performance.

This study seeks to address the empirical gap in literature regarding the link between secondary education outcomes and county-level unemployment rates in Romania in 2023.

2. Literature review

2.1. Socioeconomic Status and Educational Achievement

A substantial body of research highlights the significant influence of socioeconomic status (SES) on students' academic performance, both in national and international contexts. Common proxies for SES include parental education and household income, which have been shown to moderately correlate with academic outcomes [1,2]. In Sweden, Boman found that regional disparities in educational performance among secondary students could be partially explained by factors such as migration background and the formal competence of teachers [3,4]. Similarly, Daniele emphasized that relative poverty, acting as a multidimensional indicator, contributes to regional inequalities in academic achievement, reinforcing the notion of a socioeconomic gradient in education [5].

2.2. Local Economic Conditions and Education

Education also plays a pivotal role in shaping regional development. Empirical studies suggest that university presence and regional orientation positively influence local socioeconomic outcomes. Mavri et al.

found that educational institutions contribute to economic well-being at the local level [6], while Weßling and Bechler noted that higher regional unemployment rates may paradoxically increase university enrollment rates, as individuals seek to improve their employment prospects through education [7].

2.3. Educational Attainment and Labor Market Outcomes

There is a well-documented link between educational attainment and labor market success. According to the OECD [8], employment rates increase significantly with the level of education attained; over 80% of individuals with tertiary education are employed, compared to just over 70% for those with upper secondary education and under 60% for those without it. Furthermore, Riddell and Song showed that individuals with higher levels of formal education are less likely to experience unemployment and more likely to achieve re-employment after job loss [9].

A study by Barbu found that fluctuations in national unemployment rates affect higher education outcomes such as enrollment and retention, with effects varying by gender, ethnicity, and SES [10]. This indicates that labor market conditions can influence educational decisions and trajectories, reinforcing a bidirectional relationship.

2.4. Educational Attainment and Labor Market Outcomes

Beyond individual outcomes, higher educational attainment yields broader community-level benefits. The Alliance for Excellent Education showed that increased high school graduation rates lead to measurable economic gains for communities, benefiting various stakeholders from local business owners to financial institutions [11]. These findings support the notion that educational success is not only a personal asset but also a driver of collective economic development.

3. Methodology

This study relies on secondary statistical

data collected for all 42 counties of Romania for the year 2023. The Baccalaureate pass rates were obtained from official national sources, while unemployment rates were sourced from the Ministry of Labor's official statistics page [12]. Data on average net monthly income per county were retrieved from county-level statistics website [13]. In addition, Durbin-Watson significance tables were consulted via the University of Notre Dame repository [14].

For data processing and analysis, Microsoft Excel and SPSS were employed to organize datasets, compute descriptive statistics, and apply classical econometric methods. GeoDa software was used for conducting spatial regression analysis.

The analytical approach consisted of several stages. Bivariate correlation analyses were performed using both Pearson and Spearman correlation coefficients to assess the linear and monotonic associations, respectively, between the Baccalaureate pass rate and the unemployment rate. A multiple linear regression model was specified to evaluate the joint predictive power of Baccalaureate pass rates and average income levels on county-level unemployment rates.

The equation for the regression model can be written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (1)$$

where:

Y is the dependent variable (Unemployment rate);

X_1 is the first predictor variable (Baccalaureate Promovability rate);

X_2 is the second predictor variable (Average income);

β_0 is the intercept (constant term);

β_1 is the coefficient for X_1 ;

β_2 is the coefficient for X_2 ;

ε is the error term.

Given the geographical nature of the dataset (county-level observations), spatial econometric analysis was conducted to detect and control for potential spatial dependence. Specifically, a Spatial Lag

Model (SAR) was implemented using GeoDa, allowing the model to account for the influence of neighboring counties' unemployment rates on each individual observation.

4. Results and discussions

4.1. Pearson and Spearman Correlations

There are 42 counties in Romania, resulting in a sample size of $N = 42$. Both the Baccalaureate pass rate (graduation rate) and the unemployment rate are expressed as percentages. The standard deviation of the unemployment rate was approximately 1.84 percentage points, which, when compared to its mean value of 3.64%, represents a coefficient of variation of about 50.5%. This implies that although the absolute values of unemployment rates appear numerically low, the variability across counties is quite substantial in proportional terms. Such a high coefficient suggests marked regional disparities in labor market conditions, with some counties experiencing very low unemployment and others significantly higher rates relative to the national average. In contrast, the Baccalaureate pass rate exhibited a standard deviation of 7.43 percentage points around a mean of 72.87%. This results in a coefficient of variation of roughly 10.2%, which, although lower in relative terms compared to unemployment, reflects considerable absolute variation in educational achievement across regions. The larger spread in educational outcomes may reflect systemic differences such as disparities in school funding or teacher quality, rural vs. urban divides in educational infrastructure, or differing socioeconomic and cultural conditions that affect student performance. These contrasting patterns suggest that while unemployment rates vary sharply in proportion to their mean (possibly due to localized economic vulnerabilities or regional industrial decline), educational performance varies more broadly in absolute terms, pointing to structural inequalities in Romania's education system.

Table 1 Descriptive statistics for county-level indicators (2023)

	Unemployment rate	Graduation rate
N	42	42
Mean	3,6388	72,8667%
Std. Deviation	1,83669	7,43379%

The Pearson correlation coefficient (Table 2) between the Baccalaureate pass rate and the unemployment rate is -0.218, indicating a negative linear relationship: as pass rates increase, unemployment rates tend to decrease. However, this correlation is weak in magnitude and not statistically significant at the conventional 0.05 level ($p = 0.165$), given the sample size of $N = 42$ counties. This suggests that, while there is a slight inverse association between educational success and unemployment, the relationship is not strong enough to rule out the possibility that it occurred by chance in this sample.

In addition to the Pearson correlation, the Spearman rank-order correlation was computed to assess the strength and direction of the monotonic relationship between the Baccalaureate pass rate and the unemployment rate. As shown in Table 2, the Spearman coefficient ($\rho = -0.241$) also indicates a weak to moderate negative association: higher pass rates tend to correspond to lower unemployment rates, and vice versa. However, this correlation is likewise not statistically significant at the 0.05 level ($p = 0.125$). The lack of statistical significance in both tests implies that the observed associations may not be strong or robust enough to draw definitive conclusions based on this dataset alone. This may be due to sample size limitations, measurement variability, or the influence of unaccounted confounding factors such as labor mobility, informal employment, or regional economic structures.

4.2. Multiple Regression Model

To further explore the combined effects of Baccalaureate pass rates and average income levels on the county-level unemployment rate, a multiple linear

regression model was estimated. The model yielded an overall coefficient of determination (R^2) of 0.274, indicating that approximately 27.4% of the variance in the unemployment rate is jointly explained by the two predictors. The adjusted R^2 of 0.237 suggests a modest explanatory power, accounting for the number of predictors relative to the sample size.

The overall model was statistically significant, as evidenced by the F-statistic ($F = 7.37$, $p = 0.002$). This supports the rejection of the null hypothesis that all regression coefficients are equal to zero, implying that at least one of the predictors significantly contributes to the model. Upon examining the individual regression coefficients, important distinctions emerge:

- The coefficient for the Baccalaureate pass rate was negative ($\beta = -0.025$), aligning with the hypothesis that higher educational achievement may reduce unemployment. However, this effect was not statistically significant ($p = 0.469$), indicating that, after controlling for income, educational performance alone does not exert a meaningful influence on county-level unemployment variability.
- In contrast, the average income variable exhibited a statistically significant negative effect ($\beta = -0.002$, $p = 0.001$), suggesting that higher income levels are robustly associated with lower unemployment rates. This relationship is both substantively and statistically meaningful, highlighting income as the more powerful explanatory factor in the model.

Together, these results suggest that while educational achievement and economic well-being are correlated, income emerges as the dominant driver of regional

unemployment disparities.

The analysis (Table 2) reveals that the Baccalaureate pass rate does not emerge as a statistically significant predictor of county-level unemployment rates ($p = 0.469$). This suggests that, when controlling for income, the direct impact of educational achievement on employment outcomes may be limited or mediated by other structural factors, such as job availability, migration, or labor market mismatches. In contrast, average net income demonstrates a robust and statistically significant negative

association with unemployment ($p = 0.001$), underscoring its central role in shaping labor market dynamics. This implies that counties with higher income levels tend to experience systematically lower unemployment, likely due to stronger local economies and better job retention mechanisms. Thus, while educational attainment is important, this model suggests that economic vitality is a more powerful driver of employment outcomes in the Romanian context.

Table 2 Multiple Regression Coefficients Table

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	11.802	2.750		4.291	0.000
Graduation rate	-0.025	0.035	-0.103	-0.731	0.469
Average income	-0.002	0.000	-0.490	-3.489	0.001

Note: Dependent variable = Unemployment

Based on Equation (1), the initial estimation model at the county level can be described as follows:

$$Unemployment_c = \beta_0 + \beta_1 X_c^{Promovability} + \beta_2 X_c^{Income} + \varepsilon_c$$

where:

$Unemployment_c$ is the percentage of the rate of unemployment in each county c ;

$\beta_1 X_c^{Promovability}$ are the coefficients of graduation rate in each county;

$\beta_2 X_c^{Income}$ are the coefficients of average income levels in each county;

ε_c is an error term.

Based on the coefficients provided in Table 3, the equation becomes:

$$Y = 11.802 - 0.025 \cdot X_1 + 0.002 \cdot X_2 + \varepsilon$$

To assess the presence of autocorrelation in the residuals of the multiple linear regression model, the Durbin-Watson (DW) test was conducted (Table 3). The resulting test statistic was $DW = 2.190$, which falls between the upper critical value ($dU = 1.600$) and 4 minus the lower critical value ($4 - dL = 2.609$). According to Durbin-Watson significance thresholds, when the

DW statistic lies within the interval ($dU, 4 - dL$), the test is inconclusive for positive or negative autocorrelation but does not provide sufficient evidence to suspect autocorrelation in the residuals.

Thus, the residuals from the regression model can be reasonably assumed to be uncorrelated, satisfying one of the key assumptions of the classical linear regression model. This enhances the reliability of the coefficient estimates and supports the statistical validity of the model's inference.

Table 3 Durbin-Watson Test results

dL	1.391
dU	1.600
$4 - dL$	2.609
$4 - dU$	2.400
DW	2.190

Figure 1 depicts a scatterplot illustrating the relationship between the Baccalaureate pass rate (horizontal axis) and the unemployment rate (vertical axis) across

Romanian counties. A negative linear trend is visible, suggesting that counties with higher educational achievement tend to exhibit lower unemployment levels. This is consistent with theoretical expectations linking human capital accumulation to improved labor market outcomes.

The regression equation shown on the plot quantifies this inverse association. However, the coefficient of determination ($R^2 = 0.048$) reveals that the linear model explains only about 4.8% of the variance in unemployment rates. This indicates a very weak explanatory power, and the large dispersion of points around the regression line suggests that the relationship lacks precision and may be influenced by other unaccounted variables.

Notably, a clear outlier, marked in red in the lower-left corner of the graph, represents a county with an exceptionally low unemployment rate and a relatively low pass rate. This deviates markedly from the general pattern, potentially indicating the influence of localized factors, such as dominant industries, external labor migration, or temporary employment programs, that disrupt the broader national trend.

In summary, while the scatterplot supports the hypothesized negative association between education and unemployment, the weak linear fit and presence of outliers point to a more complex interplay of regional dynamics that may not be captured by simple bivariate analysis.

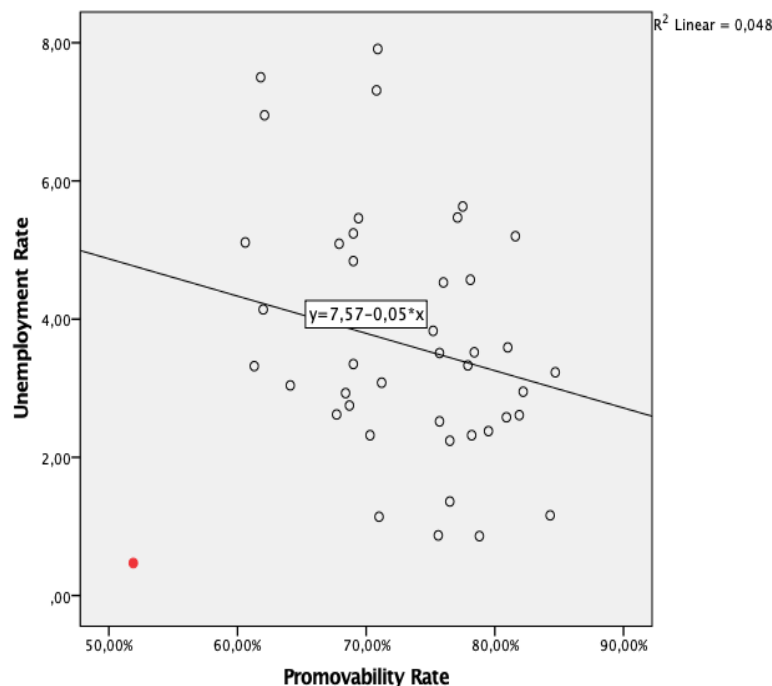


Figure 1: Scatterplot showing the relationship between Unemployment Rate and Graduation Rate

4.3. Spatial Lag Model

A Spatial Lag Model was estimated using GeoDa to account for spatial dependencies in unemployment across Romanian counties. The spatial lag coefficient ($\rho = 0.905, p < 0.001$) indicates strong positive spatial autocorrelation, meaning that counties with high unemployment are often surrounded by others with similar rates.

The model explains 24.2% of the variation in unemployment ($R^2 = 0.242$), comparable to the non-spatial model. Although this represents a modest level of explanatory power, it is comparable to the non-spatial model and confirms that space matters in understanding unemployment dynamics. However, the Likelihood Ratio test ($p = 0.231$) suggests the spatial model does not significantly improve overall fit,

though it captures inter-county dependencies well.

The results of the spatial lag regression model highlight the influence of both spatial and economic factors on unemployment across Romanian counties (Table 4). The spatially lagged unemployment variable exhibits a strong and statistically significant positive effect ($\rho = 0.905, p < 0.001$), indicating that unemployment rates in one county are closely associated with those in neighboring counties. This confirms the presence of spatial dependence and suggests that labor market challenges are regional rather than isolated. Average income also emerges as a robust and significant predictor ($\beta =$

$-0.00166, p < 0.001$), with higher income levels associated with lower unemployment rates, reinforcing the central role of economic conditions in shaping labor market outcomes. In contrast, the Baccalaureate pass rate shows a negative but statistically non-significant effect ($\beta = -2.59, p = 0.448$), suggesting that once spatial dynamics and income are accounted for, educational attainment alone does not significantly influence unemployment at the county level. These findings underscore the importance of incorporating spatial context and economic strength into policy strategies aimed at reducing regional unemployment disparities.

Table 4 Spatial Lag Regression results

Variable	Coefficient	Std.Error	z-value	Probability
W_UnempRate	0.904762	0.0680053	13.3043	0.00000
Constant	8.68957	2.71888	3.19602	0.00139
PromovRate	-2.5916	3.41636	-0.758586	0.44810
Income	-0.00166269	0.000459255	-3.6204	0.00029

The model suggests that regional income differences and spatial spillovers are more influential than education alone in shaping unemployment patterns.

The next equation represents the relationship between the unemployment rate and the promovability rate, average income, and spatially lagged unemployment rate, accounting for the coefficients provided in the Table 4:

$$\begin{aligned} UnempRate = & 0.904762 \cdot W_UnempRate \\ & - 2.5916 \cdot PromovRate \\ & - 0.00166269 \cdot Income \\ & + 8.68957 + \epsilon \end{aligned}$$

5. Conclusions

This study explored the complex interrelationships between Baccalaureate pass rates, unemployment rates, and average income levels across Romanian counties. The statistical analysis revealed a weak to moderate negative correlation between educational attainment and unemployment; however, this association did not reach statistical significance,

suggesting limited predictive power of pass rates alone. In contrast, average income emerged as a significant and consistent predictor, emphasizing the central role of economic conditions in shaping regional labor market outcomes.

The spatial regression analysis provided additional insight by confirming the existence of spatial autocorrelation in unemployment rates. Counties with high unemployment were frequently clustered near others with similarly high rates, reflecting regional labor market interdependencies. This spatial dimension highlights the importance of place-based policies that account for geographic disparities in economic development and employment opportunities.

Despite its contributions, the study is not without limitations. The use of secondary data and the relatively small sample size ($N = 42$ counties) may restrict the generalizability of the findings. Moreover, the analysis did not account for other potentially influential variables, such as

sectoral employment structure, migration patterns, or infrastructure, which may moderate or mediate the observed relationships.

Nevertheless, the findings offer guidance for policymakers. Efforts to reduce unemployment and enhance labor market resilience should address both economic inequalities and spatial clustering effects, alongside educational initiatives. Future research should aim to develop more

comprehensive models by incorporating additional regional indicators and longitudinal data to better capture the dynamic and multidimensional nature of unemployment in Romania.

Acknowledgements

This paper was co-financed by The Bucharest University of Economic Studies (ASE) during the PhD programme.

References

- [1] Kim S won, Cho H, Kim LY. Socioeconomic Status and Academic Outcomes in Developing Countries: A Meta-Analysis. *Review of Educational Research*. 2019 Sep 25; 89(6):875–916.
- [2] Gustafsson JE, Yang Hansen K. Changes in the Impact of Family Education on Student Educational Achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research*. 2017 Apr 5; 62(5):719–36.
- [3] Boman B. Regional Differences in Educational Achievement among Swedish Grade 9 Students. *Scandinavian Journal of Educational Research*. 2021 Mar 18; 66:1–16.
- [4] Boman B. Regional differences in educational achievement: A replication study of municipality data. *Frontiers in Education*. 2022 Jul 28; 7.
- [5] Daniele V. Socioeconomic inequality and regional disparities in educational achievement: The role of relative poverty. *Intelligence*. 2021 Jan; 84:101515.
- [6] Mavri M, Angelis V, Dimaki K. The Impact of Educational Activities on Regional Development. *RePEc: Research Papers in Economics*. 2011 Sep 1.
- [7] Weßling K, Bechler N. Where do regional influences matter? The impact of socio-spatial indicators on transitions from secondary school to university. *Review of Regional Research*. 2019 Oct 1;39(2):163–88.
- [8] OECD. How does education affect employment rates? In: *Education at a Glance 2014: Highlights*. Paris: OECD Publishing; 2014.
- [9] Riddell WC, Song X. The impact of education on unemployment incidence and re-employment success: Evidence from the U.S. labour market. *Labour Economics*. 2011 Aug;18(4):453–63.
- [10] Barbu DG. Relationship between unemployment and college enrollment and success outcomes [Doctoral Dissertation]. [The Florida State University]; 2015.
- [11] Alliance for Excellent Education. The Economic Benefits from Halving the Dropout Rate: A Boom to Businesses in the Nation’s Largest Metropolitan Areas. 2010 Jan.
- [12] Ministry of Labour and Social Solidarity. *Statistici-șomaj 2023*. Available from: <https://mmuncii.ro/j33/index.php/ro/transparenta/statistici/datestatistice?id=6505:statistici-somaj-2023&catid=29:domenii>. Accessed Jun 2024.
- [13] Statistici Romania, [Statisticiromania.ro](http://statisticiromania.ro) 2024. Available from: <http://statisticiromania.ro/statistici-judetene>. Accessed on Jun, 2024.
- [14] University of Notre Dame. A Durbin-Watson Significance Tables. Available from: https://www3.nd.edu/~wevans1/econ30331/Durbin_Watson_tables.pdf. Accessed Jun 2024.