HOW IMPORTANT EDUCATION IS? AN EXPLORATIVE ANALYSIS OF THE DROPOUT PHENOMENON IN EASTERN EUROPE

Piera CASCIOLI[1], Emiliano COLANTONIO[1], Daniel BADULESCU[2]

[1]Department of Philosophical, Pedagogical and Economic-Quantitative Sciences, University of Chieti-Pescara, Pescara, Italy
[2]Department of Economics and Business, Faculty of Economic Sciences, University of Oradea, Oradea, Romania

piera.cascioli@unich.it
colantonio@unich.it
dbadulescu@uoradea.ro

Abstract: *Early school leaving is a structural problem related to the socioeconomic environment involving aspects of local culture and customs that characterize a country. If not controlled, with the intent of hindering it, this phenomenon can generate unemployment, social exclusion, poverty, health problems and a slowdown in the country's economic growth. Since education policies are one of the most important concerns of national governments and international organizations, the UN 2030 Agenda, containing the common ambitions of the signatory countries, proposes 17 sustainable development goals and specifically, among these, inclusive and equitable quality education. In some Eastern European countries, the percentage of young people who decide not to continue their studies is really too high, exceeding 10%. Therefore, the aim of the paper is to study the dropout phenomenon in this context. We employ a panel vector autoregressive model in first differences to test complex dynamic relationships between share of the population aged 18 to 24 not involved in any education or training (as a proxy for the dropout phenomenon), GDP per capita (as a proxy for a country's wealth and industrial modernity), gross domestic expenditure on R&D in higher education (as a proxy for a country's effort to improve its education system), and share of people reporting crime, violence or vandalism (as a proxy for a country's social condition). The study concerns 10 Eastern European countries for the period 2000-2021. The results show that the government expenditure in education is negatively related to the dropout rate. Moreover, the increase in the relative number of early school leavers seems to stimulate a worsening in social conditions, with an increase in cases of crime, violence and vandalism. Finally, both an increase in dropouts and a deterioration in social conditions generate negative effects on the well-being of the community and economic growth. Improving the quality of the education system is therefore crucial (this is also confirmed by the variance decomposition analysis), even if it may not be enough: from the impulse response functions analysis, indeed, a shock exerted on the government expenditure on R&D in higher education produces positive effects on the* *(declining) dropout rate, but only for a short period. Policymakers should therefore make constant efforts to reduce the early school leaving.*

Keywords: *school dropouts; education; human capital; socioeconomic development*

JEL Classification: *I2; H52*.

1. Introduction

The importance that a society gives to education is the basis for human capital development. Education is therefore an important driver for economic growth (Coman et al., 2023) and a key factor for the knowledge-based economy (Castagna et al., 2010). The education system can be one of the fundamental pillars for the socioeconomic development of a country. Hence, while it is true that education is one of the factors that can ensure economic growth, the interruption of studies is a real obstacle to socioeconomic development. Specifically, early school leaving is a structural problem related to the socioeconomic context involving aspects of local culture and customs that characterize a country (Colombo 2010).
Among the reasons for dropping out there are the medium- and low-tech production contexts (European Commission, 2015), the poor quality of the school system (Archambault et al., 2009), the limited economic possibilities of the family (Bynner and Parsons, 2002), the low importance that families attribute to the educational development of their offspring (Gorard, 2010), the poor integration with peers (Wehlage et al., 1989), as well as the presence of delinquency in the area of residence (Ripamonti and Barberis, 2018).
Education policies are among the most important concerns of national governments and international organizations. The UN 2030 Agenda, containing the common ambitions of the signatory countries, proposes 17 sustainable development goals and, among these, inclusive and equitable quality education. In the European context, cutting early school leaving is a priority. Already in 2010, recording an average school dropout rate of 16.4%, the European Commission included the reduction of the phenomenon among the objectives of the Europe 2020 strategy. On average, the school dropout rate has decreased from 13.4% in 2011 to 10.2% in 2019 across Europe. Although some progress has been made, the fight against early school leaving cannot stop, especially considering that in some countries such as Germany, Spain, Italy, Portugal, Romania, Bulgaria, and Hungary, rates are still too high (Grosseck, et al., 2020; Caroleo et al., 2020; Palmisano et al., 2022).

The purpose of the work is to study the phenomenon of early school leaving in 10 Eastern European countries, investigating the factors that can influence the choice of young people not to continue with their education, as well as the related consequences. We therefore employ a panel vector autoregressive model in first differences to test complex dynamic relationships between dropout phenomenon (proxied by the share of the total population aged 18 to 24 not involved in any education or training, DROP), wealth and industrial modernity of a country (proxied by the GDP per capita, GDP), government effort to improve the education system (proxied by the gross domestic expenditure on R&D in higher education, EXP), and social conditions (proxied by the share of people reporting crime, violence or vandalism in the area of residence, SOC).

2. Empirical analysis

2.1. Methodology and model specification

The aim of the analysis is to highlight the role that some key factors can play in determining the dropout phenomenon, as well as its possible consequences, in a panel of 10 Eastern European countries (Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovak Republic, and Slovenia), using the longer time span possible from 2000 to 2021. We employ a PVAR model based on a system of equations in which all the variables are treated as endogenous. Such a feature helps in exploring multiple relationships between variables. Moreover, the PVAR allows to capture the effects of one exogenous shock in one variable to another variable in the system, while keeping all other variables invariant. This permits to highlight bidirectional dynamic effects and potential path dependences. In line with the recent literature, we propose a PVAR model based on the indicators listed in Table 1.

Table 1: Data description and sources

Variable	Definition	Source
DROP	Share of the total population aged 18 to 24 who has completed at most lower secondary education and is not involved in further education or training	Eurostat
GDP	GDP per capita (constant 2015 US$)	World Bank
EXP	Government expenditure on R&D in the higher education sector (%, relative to GDP)	Eurostat
SOC	Share of people reporting crime, violence or vandalism in their local area	Eurostat

The conventional PVAR model is given by the following system of equations (Love and Zicchino, 2006):

$$X_{it} = f_i + g_t + \square(L)X_{it} + \square_{it} \quad (1)$$

where X_{it} is the vector of stationary variables in our analysis, f_i is a vector of country fixed effects, g_t is a vector of year fixed effects, $\square(L)X_{it}$ is a square matrix of polynomials in the lag operator, and $\square_{it}$ is the random error term (later, d denotes the first difference operator). Table 2 shows the main descriptive statistics.

Table 2: Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
DROP	9.54	4.53	3.00	22.40
GDP	12916.46	4704.08	3717.90	24744.84
EXP	0.25	0.16	0.04	0.73
SOC	12.23	6.71	3.20	28.70

2.2. Empirical testing

The use of the PVAR model requires stationary variables, since non-stationarity can cause spurious results in the context of VAR and panel analyses. Hence, the first step of the analysis is to test for the stationarity of the various series through first- and second-generation unit root tests. Specifically, two first-generation unit root tests (IPS and MW) and one second-generation unit root test (Pesaran) were performed. All tests are characterized by a null hypothesis assuming a unit root. The results reported in Table 3 show that the variables are non-stationary in level.

Table 3: Unit root tests: variables in level

Variable	IPS W-t-bar	MW	Pesaran
DROP	-2.9572	19.306	0.475
GDP	3.2615	18.995	0.083*
EXP	-0.8227	23.075	0.088*
SOC	-0.961	26.411	0.512

Note: $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

In these cases, a possible solution is the use of the first-difference transformation (Gyimah et al, 2022; Acheampong, 2018). The above listed unit root tests were also performed to the transformed variables. Based on the results listed in Table 4, we can conclude that all the series are integrated of order one (I(1)), i.e. all the chosen variables are stationary after the first difference transformation.

Table 4: Unit root tests: variables in first differences

Variable	IPS W-t-bar	MW	Pesaran
dREN	-7.426***	72.602***	-4.352***
dGDP	-8.787***	79.066***	-1.670**
dRAD	-10.684***	105.252***	-3.915***
dWOM	-6.619***	51.970***	-3.422***

Note: *p < 0.1; **p < 0.05; ***p < 0.01

Subsequently, four cointegration tests introduced by Westerlund (2007) were performed to check for possible cross-section interdependence. These tests assume the null hypothesis of no cointegration, which cannot be rejected based on the results of all four tests listed in Table 5. The choice of the first-difference estimates is therefore supported, since the variables in level are non-cointegrated, as well as non-stationary.

Table 5: Cointegration tests

Statistic	Value	p-value
G□	-1.66	0.38
G□	-4.12	0.68
P□	-4.42	0.38
P□	-2.52	0.70

Note: p-value are robust critical values obtained through bootstrapping with 100 replications

We also examined the correlation matrix and variance inflation factor (VIF) in order to assess whether collinearity and multicollinearity were a problem for our analysis. Given the low correlation values and low mean VIF and VIF reported in Table 6 (dDROP has been used as the dependent variable), we can conclude that collinearity and multicollinearity are not a concern.

Table 6: Correlation matrices and VIF statistics

	dDROP	dGDP	dEXP	dSOC
dDROP	1.00			
dGDP	0.01	1.00		
dEXP	0.07	0.01	1.00	
dSOC	0.14	-0.16	0.04	1.00
VIF		1.03	1.00	1.03
mean VIF	1.02			

The selection of the optimal lag is the last preliminary step of our analysis. According to the econometric literature, the commonly used criteria are the Moment Akaike Information Criterion (MAIC), the Moment Bayesian Information Criterion (MBIC), and the Moment Hannan-Quinn Information Criterion (MHQIC), which are maximum likelihood-based selection criteria. Following Andrews and Lu (2001), the ideal lag length should minimize the moment model selection criteria MBIC, MAIC, and MHQIC. Accordingly, considering the results shown in Table 7, the optimal model is a first order PVAR.

Table 7: Lag order selection criteria

Lag	MBIC	MAIC	MHQIC
1	-199.184	-51.575	-111.514
2	-136.335	-37.929	-77.889
3	-76.633	-27.430	-47.410

We removed the country fixed effects f_i in Eq. (1) by means of the first difference transformation. This method, however, may generate the so-called Nickell bias (1981) due to the correlation between the first-differenced lag and the first-differenced error term, both dependent on □$_{it-1}$. In this context, estimating the model using OLS will produce biased and inconsistent results (Baltagi, 2008). We therefore used forward mean-differencing, also referred to as the Helmert transformation (Ht) (Arellano and Bover, 1995; Love and Zicchino, 2006) to overcome this problem. Moreover, applying the Ht to data produces the same result as applying the Ht to demeaned data, so that the year fixed effects g_t in Eq. (1) are also removed. Therefore, the system can be estimated using the Generalized Method of Moments and the lagged values of regressors can be used as instruments.

2.3. Results

The first order PVAR results are shown in Table 8.

Table 8: PVAR results

		Dependent variables			
		dDROP	**dGDP**	**dEXP**	**dSOC**
Lagged independent variables	**dDROP**	0.354***	-225.804***	-0.049*	0.659***
	dGDP	0.000	0.018	0.000	0.001**
	dEXP	-1.846*	-3651.894***	-0.016	3.112
	dSOC	-0.046	-206.323***	0.006	0.246***

Note: *p < 0.1; **p < 0.05; ***p < 0.01

In line with the literature (Dincă, 2019), an increase in EXP generates a decrease in DROP: likely, investments in education, improvements in the quality of the school system (Archambault et al. 2009), as well as the teacher training (Dupper, 1993), could help to increase young people's desire for knowledge and learning. Unexpectedly GDP and SOC do not seem to have any statistically significant impact on DROP, at least in the short run. Based on the results, DROP shows a path dependence, which can have a double interpretation: the absence of programs to reintegrate young dropouts into the educational sector usually generate a negative spiral over time; on the other hand, an attenuation of the dropout phenomenon can persist over the years.

Our empirical analysis also shows that DROP and SOC have a statistically significant negative impact on GDP. In line with previous studies, the lack of adequate education and delinquent behaviour limit job opportunities (especially for young people), with negative repercussions on the wealth and well-being of the community (Freeman, 1991).

The results also highlight that an increase in EXP generates a reduction in GDP. Likely, investments in education are not productive, at least in the short run. The resources allocated to the education sector are destined to generate their (positive) effects over the years. Moreover, it is possible to underline the existence of an indirect relationship between the variables involved: as seen before, an increase in EXP generates a reduction in DROP, with a subsequent positive influence on GDP.

In line with the literature (Dennison, 2022), an increase in DROP induces an increase in SOC, that is criminal involvement is associated with early school leaving; in other words, the dropout phenomenon can stimulate a worsening in social conditions, with an increase in cases of crime, violence and vandalism.

Following Lutkepohl (2005), the stability of the PVAR model was verified since the eigenvalues are strictly less than 1 (see Table 9). Stability implies that the PVAR model is invertible, thus allowing bidirectional interpretations, as well as impulse-response functions estimates and the variance decomposition analysis (Abrigo and Love, 2016). In addition, the over-identification restriction test (Hansen's J chi2) is equal to 46.89 ($p = 0.518$): this confirms the goodness-of-fit of the model, since the null hypothesis that the over-identification restrictions are valid is verified (i.e. instrumental variables included are valid instruments and uncorrelated with the error term, while instruments not included are correctly excluded).

Table 9: Eigenvalue stability condition

Real	Imaginary	Modulus
0.446	0.000	0.446
0.138	0.371	0.395
0.138	-0.371	0.395
-0.121	0.000	0.121

Furthermore, the Granger causality tests, which investigates the null hypothesis of absence of causality, reveal bi-directional causality and confirm the presence of endogeneity (see Table 10).

Table 10: Granger causality test

Equation Variable	Excluded Variables	Chi2	p-value
dDROP	dGDP	1.921	0.166
	dEXP	3.038	0.081
	dSOC	1.752	0.186
	ALL	6.836	0.077
dGDP	dDROP	10.164	0.001
	dEXP	8.840	0.003
	dSOC	11.267	0.001
	ALL	28.788	0.000
dEXP	dDROP	10.453	0.001
	dGDP	0.836	0.361
	dSOC	1.885	0.170
	ALL	11.459	0.009
dSOC	dDROP	15.491	0.000
	dGDP	4.483	0.034
	dEXP	1.749	0.186
	ALL	21.013	0.000

Table 11 shows the variance decomposition (following the Cholesky decomposition using 1000 Monte Carlo simulations for 10 periods), which evaluates the percentage change in a variable which is explained by the shock to another variable over time. The results show that each variable is mainly affected by its lag. Specifically, DROP is mainly stimulated by EXP, while GDP is strongly influenced by DROP and SOC, and finally SOC is heavily stimulated by DROP.

Table 11: Variance decomposition analysis

		Impulse Variable			
		dDROP	**dGDP**	**dEXP**	**dSOC**
Response variable	**dDROP**	94.76%	1.51%	1.72%	2.01%
	dGDP	7.44%	71.39%	5.19%	15.98%
	dEXP	29.07%	0.79%	68.31%	1.84%
	dSOC	10.71%	9.13%	1.34%	78.82%

Note: Variation in response variable explained by the impulse variables in the columns (10 periods ahead)

Each impulse response function in Figure 1 provides a picture of the reaction of one variable to one standard deviation shock on another variable, all other shocks being equal to zero (a Gaussian approximation based on 200 Monte Carlo simulations was used to estimate the impulse response functions, again following the Cholesky decomposition). When a shock is exerted on a variable in the current period, the response variable usually shows a remarkable response during the first years, followed by a slight fluctuation thereafter. Specifically, a shock in education spending reduces dropouts during the early periods; the effect, however, decreases quickly, indicating that the government effort should be constant to have a long-lasting reduction in the dropout rate. Moreover, positive shocks in DROP and SOC lead to a reduction in wealth and well-being, causing, among other things, a slowdown in the economy.

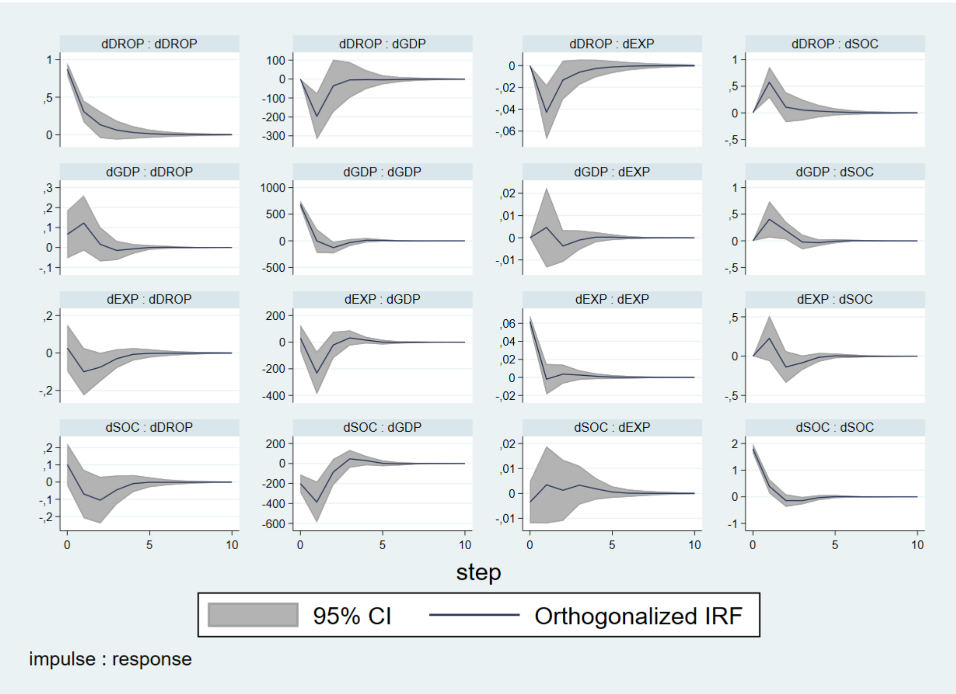

Figure 1: Impulse Response Analysis

3. Conclusions and Policy Implications

The paper tries to contribute to the literature on the determinants of the dropout phenomenon and its consequences, focusing on the role played by proxies of human capital (DROP), social capital (SOC), wealth and modernity (GDP), and expenditure in education (EXP). Specifically, we employed a PVAR model in first differences to test the relationships among the aforementioned variables, during the period 2000-2021, in 10 Eastern European countries, where the early school leaving is still a problem.

The results highlight that DROP has a path dependence over time; specifically, the absence of programs to reintegrate young dropouts into the educational sector can generate a negative spiral. Policy makers should therefore dedicate a constant effort to the attenuation of the dropout phenomenon, preventing it from persisting over the years. Specifically, one of the main drivers of DROP appears to be spending on education. Adequate resources, both financial and non (support programs, qualified personnel, etc.), should therefore be allocated in the school system to ensure full integration of young people, especially in cases of problems (disability, economic, etc.) that can limit the scholastic career.

Counterintuitively, GDP seem to have no statistically significant impact on DROP, at least in the short run. On the other hand, DROP negatively influences GDP: likely, early school leaving limits job opportunities, with negative consequences on the economic growth of a country. Again, it is therefore important that policy makers invest in education with the aim of drastically reducing the dropout phenomenon.

Finally, an increase in the dropout rate also generates an in increase in criminal involvement, with further negative consequences on the wealth and well-being of the community.

Government investments in the higher education sector are essential to tackle the dropout phenomenon, even if they negatively impact GDP in the short run. These investments, indeed, could not be immediately productive. The resources allocated to the education sector will generate their (positive) effects over the years. Furthermore, the existence of an indirect relationship between the variables involved has been highlight: an increase in EXP tends to reduce DROP, with a subsequent positive impact on GDP.

References

1. Abrigo, M. R. M. and Love, I. (2016) “Estimation of panel vector autoregression in Stata”, *Stata Journal*, Vol. 16, No. 3, pp. 778-804.

2. Acheampong, A. O. (2018) “Economic growth, CO2 emissions and energy consumption: What causes what and where?”, *Energy Economics*, Vol. 74, pp. 677-692.

3. Andrews, D. W. K. and Lu, B. (2001) "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models", *Journal of Econometrics*, 101, Vol. 1, pp. 123-164.
4. Archambault, I., Janosz, M., Morizot, J. and Pagani, L. (2009) "Adolescent behavioral, affective andcognitive engagement in school: Relationship to dropout", *Journal of School Health*, Vol. 79, No. 9, pp. 408-415.
5. Arellano, M. and Bover, O. (1995) "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, Vol. 68, pp. 29-51.
6. Baltagi, B. H. (2008) Econometric analysis of panel data, Wiley, Chichester.
7. Bynner, J., and Parsons, S. (2002) "Social exclusion and the transition from school to work: The case ofyoung people not in education, employment, or training (NEET)", *Journal of Vocational Behaviour*, Vol. 60, No. 2, pp. 289-309.
8. Caroleo, F. E., Rocca, A., Mazzocchi, P., and Quintano, C. (2020) "Being NEET in Europe before and after the economic crisis: An analysis of the micro and macro determinants", *Social Indicators Research*, Vol. 149, No. 3, pp. 991-1024.
9. Castagna, A., Colantonio, E., Furia, D., and Mattoscio, N. (2010) "Does education play a relevant role inglobalization?" *Procedia - Social and Behavioral Sciences*, Vol. 2, No. 2, pp. 3742-3750.
10. Colombo, M. (2010) Dispersione scolastica e politiche per il successo formativo: dalla ricerca sugli early school leaver alle proposte di innovazione, Edizioni Erickson.
11. Coman, A. C., Lupu, D., and Nuţă, F. M. (2023) "The impact of public education spending on economic growth in Central and Eastern Europe. An ARDL approach with structural break", *Economic Research-Ekonomska Istraživanja*, Vol. 36, No. 1, pp. 1261-1278.
12. Dennison, C. R. (2022) "Dropping out of college and dropping into crime", *Justice Quarterly*, Vol. 39, No. 3, 585-611.
Dincă, G. (2019) "Investment in education", Bulletin of the Transilvania University of Braşov, Series V, Economic Sciences, Vol. 12, No. 61/2, pp. 79-86.
13. Dupper, D. R. (1993) "Preventing school dropouts: Guidelines for school social work practice", *Social Work in Education*, 15(3), 141–149.
14. European Commission (2015) Macroeconomic imbalances Country Report – Italy 2015, European Commission, Brussels.
15. Freeman, R. (1991) "Crime and employment of disadvantaged youth", *National Bureau of Economic Research*, Working Paper, No. 3875.
16. Gorard, S. (2010) "School experience as a potential determinant of post-compulsory participation", *Evaluation and Research in Education*, Vol. 23, No. 1, pp. 3-17.
17. Gyimah, J., Yao, X., Tachega, M.A., Hayford, I.S., and Opoku-Mensah, E. (2022) "Renewable energy consumption and economic growth: New evidence from Ghana, *Energy*, Vol. 248, 123559.
18. Grosseck, G., Holotescu, C., and Andone, D. (2020) "Open educational resources in Romania", in Current State of Open Educational Resources in the "Belt and Road" Countries, pp. 151-173.

19. Love, I., and Zicchino, L. (2006) "Financial development and dynamic investment behavior: Evidence from panel VAR", *The Quarterly Review of Economics and Finance*, Vol. 46, pp. 190-210.
20. Lutkepohl, H. (2005) New introduction to multiple time series analysis, Springer Verlag.
21. Nickell, S. (1981) "Biases in dynamic models with fixed effects", *Econometrica*, Vol. 49, pp. 1417-1426.
22. Palmisano, F., Biagi, F., and Peragine, V. (2022) "Inequality of opportunity in tertiary education: evidence from Europe". *Research in Higher Education*, Vol. 63, pp. 514–565.
23. Ripamonti, E., and Barberis, S. (2018) "The effect of cultural capital on high school dropout: An investigation in the Italian Provinces". *Social Indicators Research*, Vol. 193, No. 3, pp. 1257-1279.
24. Wehlage, G. G., Rutter, R. A., Smith, G. A., Lesko, N., and Fernandez, R.R. (1989) Reducing the risk: Schools as communities of support, Falmer Press, Philadelphia, PA.
25. Westerlund, J. (2007) "Testing for error correction in panel data", *Oxford Bulletin of Economics and Statistics*, Vol. 69, 709-748.