

The role of EU cohesion funds in Romanian labour productivity: Insights from machine learning and econometric modelling

Research Article

Adriana AnaMaria Davidescu^{1,2}, Monica Mihaela Maer Matei^{1,2}, Marina-Diana Agafitei^{1,2*},
Maria Bianca Bolboasă^{1,2}

¹ Bucharest University of Economic Studies, Bucharest, Romania

² Department of Education, Vocational Training and Labour Market, National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania

Received 20 November 2024; Accepted 12 March 2025

Abstract: This study examines the impact of European structural and investment funds (ESIF) on labour productivity across Romania's NUTS 2 regions from 2007 until 2020. This research aims to assess how ESIF investments influence productivity imbalance while also identifying key regional determinants of economic performance, including socioeconomic structure, institutional quality, and educational attainment. Utilising a hybrid methodology integrating machine learning for variable selection and econometric modelling for effect estimation, the analysis leverages Least Absolute Shrinkage and Selection Operator to pinpoint the most influential factors and fixed effects panel regression models to quantify regional impacts. The results reveal significant disparities in the effectiveness of ESIF investments across regions, with factors such as governance quality and initial levels of development moderating the productivity outcomes. Findings signify that regions with stronger institutional infrastructures and higher baseline productivity levels benefit more meaningfully from ESIF investments. This research advances the understanding of ESIF's role in promoting equitable regional development and demonstrates the utility of combining machine learning and econometric techniques in policy evaluation.

Keywords: *Regional labour productivity • European structural and investment funds • Cohesion policy impact • Econometrics • Machine learning • EU funds efficiency*

1. Introduction

The European Union has strategically utilized Structural and Investment Funds to stimulate regional progression and financial synchronization across its Member States. These funds are crucial, not only serving as financial support but also acting as key instruments for incorporating less developed regions into the wider European economic framework, thus improving the cohesion of the Union as a whole (Breidenbach et al., 2018).

European structural and investment funds (ESIF) are intended to address a wide variety of regional challenges by supporting infrastructure development, enhancing human capital, fostering innovation, and ensuring

sustainable environmental practices. Grounded in the principle of solidarity, these funds aim to establish a harmonized developmental trajectory across Europe, directing resources to the areas most in need to level the playing field and allow all regions to contribute to and benefit from the EU's dynamic internal market (Caldas et al., 2018).

While large sums have been allocated through EU structural funds towards numerous projects aiming to stimulate development and integration across Europe's regions, rigorously assessing the true effects on members' economic disparities has proven immensely problematic (Rodríguez-Pose & Novak, 2013). The effectiveness of these funds varies greatly in different European territories, influenced by factors such as local governance capacities, existing socioeconomic structures, and regional economic specializations. This variation emphasizes the need for careful policy formulation and

* E-mail: diana.agafitei@csie.ase.ro

implementation – making it a critical area of concern both for policymakers and researchers.

Beyond economic growth, ESIF also targets qualitative aspects such as sustainable development and social inclusion. This reflects the EU's overall approach to regional development. These funds are expected not only to speed the convergence of GDP per capita among regions (Vukašina et al., 2022) but are also likely to bring broader socio-economic benefits, such as increased employment opportunities, better educational outcomes, and a better environment. The multidimensional impacts of these funds highlight the complexity of the EU's regional policy and its significance for the future of European integration.

While ESIF positively influence economic growth and development over time, looking closer at their effectiveness reveals that outcomes are far from uniform. In addition, they depend on the regional conditions as well as the form of implementation used. Midelfart-Knarvik and Overman (2002), for example, found that although European Structural Funds have had a major impact on industry location, their effect has frequently been to offset the comparative advantages of the regions by drawing R&D-intensive industries to areas with low-skilled labour endowments. Similarly, research on the environmental effects of those funds points to mixed results. Incorporating environmental considerations into regional development has made some encouraging progress, while hardly any success has been achieved in sustainable development integration (Clement, 2005). In addition, Farrell (2004) has emphasized the significance of ESIF in mitigating regional inequities, pointing out that, despite growth and redistributive impacts, national, institutional, and political configurations have a key role in shaping the results. These findings suggest that although the ESIF have played a major role in promoting regional development, their effectiveness depends more on the specific policy framework and strong local governance structures that are able to exploit these resources to address the particular regional needs and problems.

This research aims to deepen the understanding of the efficacy of ESIF in fostering regional development through the lens of labour productivity within Romanian NUTS 2¹ regions from 2007 to 2020. The purpose of this study is not only to provide a comprehensive understanding of how ESIF investments influence labour productivity but also to explore the relationship between regional specificities and the success of cohesion policy investments. Thus, the research aims

to answer the following research questions: RQ1. How do ESIF influence labour productivity across Romanian NUTS 2 regions, considering the period from 2007 to 2020? RQ2. What are the key determinants of economic performance at the regional level affecting labour productivity in the context of ESIF investments? RQ3. How do regional characteristics, including socioeconomic structure, institutional framework, and initial levels of development, condition the effectiveness of ESIF in improving labour productivity?

Our article is organized as follows: In Section 2, we provide the theoretical background on the factors influencing the impact of ESIF on regional development. Following this, we describe the data used and outline the research methodology, placing it in the context of existing studies. Section 4 presents the empirical results, which are then discussed in Section 5.

2. Literature review

With Member States joining in, the European Commission makes a long-term financial plan for EU funding, aimed at boosting the economy, promoting integration and making life better for everyone in Europe (Becker et al., 2018). At the heart of this plan are the ESIF. They support the EU cohesion strategy, in pursuit of balanced economic growth with high levels of employment and effective social safety nets. ESIF is also a major source of financing for investment in many EU countries (Melecký, 2018). In addition, these funds contribute towards increased economic uniformity, towards raising environmental standards over the entire Union as well as towards the standardization of living conditions and bringing Member States together (European Commission, 2017).

Formed in the late 1950s, the European Economic Community sought to integrate markets by eliminating trade barriers, with the expectation that the unrestricted flow of goods, people, services, and capital would normalize incomes across members. Nonetheless, the European Social Fund (ESF) was established in 1958 to address disparities affecting certain worker groups (Marzinotto, 2012). By the 1970s, significant market segmentation persisted, revealing that a straightforward free-market model was inadequate for the European common market. This realization led to the formulation of EU regional policy and the European Regional Development Fund (ERDF) creation in 1975 to address these disparities further (Marzinotto, 2012). In addition to the ESF and ERDF, the ESIF includes three other funds: the European Agricultural Guidance and Guarantee Fund, established in 1962 to support the Common Agricultural Policy, the European Maritime and Fisheries Fund (EMFF) for fishing industry reforms, and the Cohesion Fund (CF) initiated in

¹ NUTS 2 (Nomenclature of Territorial Units for Statistics, level 2) is part of a standardized system used by the European Union for statistical and regional analysis.

1994 to assist the EU's poorest regions, specifically those where the Gross National Income per capita is below 90% of the EU average (Györfi *et al.*, 2016).

Given that these funds are primarily intended to boost economic growth in recipient countries and foster alignment among member states by directing financial resources towards the EU's most underdeveloped regions, there is a substantial body of research over the years exploring the effects of EU funding on economic growth using diverse research approaches. A solid body of research confirms that cohesion policy funds have significantly positive effects on regional employment, economic growth, and convergence, as demonstrated in studies by Falk and Sinabell (2008), Pellegrini *et al.* (2013), Rogge (2019), and Bachtrögl *et al.* (2020). Various studies have highlighted discrepancies in the effects of policies across diverse geographic areas, noting that less developed regions often experience different outcomes compared to more developed ones. These analyses have also contrasted lasting effects with those that diminish over the short term, indicating that some impacts may not be sustained over time (Bachtler & Gorzelak, 2007; Crescenzi & Giua, 2020). Furthermore, various analyses have shown that effects on regional employment, productivity growth, and economic convergence are either non-existent or only weakly significant. These studies also suggest that outcomes can be sensitive to the underlying assumptions in econometric specifications (Psycharis *et al.*, 2020).

In examining the macroeconomic effects of European cohesion policy, studies primarily utilize econometric approaches, with some leveraging various modelling techniques. A significant portion of this research relies on the neoclassical growth model (Barro & Sala-i-Martin, 2004), employing growth regressions (Durlauf, 2009) to analyse the impact of cohesion funds on GDP growth. Spatial econometric methods assess the indirect effects that policies have on neighbouring regions through mechanisms like trade, technological exchange, and production dynamics, often identifying positive regional spillovers from cohesion initiatives (Fidrmuc *et al.*, 2019). Another significant approach involves using quantitative general equilibrium models, which are better equipped to gauge the comprehensive impacts of regional policies (Brandsma *et al.*, 2015; Varga, 2017).

Several empirical studies have already used panel econometric methodologies, such as fixed effects (FE), random effects, and the Generalized Method of Moments (GMM), to assess the impact of EU cohesion funds on economic growth and labour productivity. Using panel FE models, Becker *et al.* (2018) found that EU regional policy had a statistically significant impact on GDP growth, with a stronger effect in regions with better institutional frameworks. Similarly, Ederveen

et al. (2006), employing FE panel models and GMM, demonstrated that the degree of effectiveness of EU structural funds depends on the quality of governance, suggesting that funds allocated to regions with weak institutions will experience a lower return in terms of economic growth. Rodríguez-Pose & Novak (2013), on the other hand, used a FE panel data model to analyse the impact of EU structural funds on economic convergence in European regions. Their analysis showed that while funds from the EU usually help the economies of its member states to grow, their effectiveness is heavily dependent on factors of context, such as the educational level, local innovation potential, and quality of governance. Furthermore, a few empirical papers highlight the effects of EU funds on labour productivity, emphasizing sectoral absorption capacity, institutional quality, and investments in human capital. Using a computable general equilibrium model, Cardenete *et al.* (2018) found that EU funds boost productivity differently across sectors, with stronger effects in regions capable of absorbing technological spillovers and infrastructure improvements. Similarly, Francesco and Pupo (2009) applied a neoclassical growth model to analyse Italian regions, showing that while EU Structural Funds reduce regional economic disparities, their impact on labour productivity remains limited due to weak institutional capacities and inefficient fund allocation. These results highlight the importance of further targeted investments in human capital and innovation to optimize long-term productivity returns from EU funding.

In conclusion, there is a vast array of studies investigating the impact of EU funds on regional development, employing a wide range of methodologies. Each methodological approach offers valuable insights, yet is accompanied by inherent limitations (Lima & Cardenete, 2008). While most studies primarily focus on the impact on economic growth, expressed as GDP growth or GDP per capita, this article makes a significant contribution by specifically analysing the effects on labour productivity. This focus helps deepen our understanding of how EU funds contribute to general economic indicators and improvements in the workforce's productivity, which is crucial for sustainable economic development.

3. Data and methodology

3.1 Data

The factors considered in estimating the impact of funds on target variables at the regional level are grouped into the following dimensions:

- **Investments:** Including variables related to cohesion policy expenditures (in current euro prices), such as the CF as a percentage of GDP, per capita European Agricultural Fund for Rural Development (EAFRD) and EMFF, ERDF per capita, ESF and Youth Employment Initiative per capita, along with other forms of investment like research and development expenditures per capita, and gross fixed capital formation per capita.
- **Socioeconomic structure:** Captured through components like education (proportion of the population with high and low education levels), economic structure (share of the main activity sector in GVA, share of knowledge-intensive services sector in GVA), accessibility (road/air), and urbanization (population density).
- **Control variables:** Considering governance quality (effects of institutional capacity), the initial level of GDP per capita (2007; convergence effects), and population size (size effects).
- **Spatial effects:** These are included through an indicator accounting for the GDP of neighbouring regions (spillover effects). This was calculated using road distances between NUTS2 regions (Persyn et al., 2020). Thus, for each region, the indicator is calculated as a weighted average of the differences between the specific region's economic growth and the economic growth recorded by the other regions.

As data sources, the model primarily uses the regional databases of Eurostat and also the Tempo database of the National Institute of Statistics. Additional data sources include the European Quality of Government Index (EQI) developed by the Quality of Government Institute at the University of Gothenburg and the regional competitiveness index of DG Regio. Cohesion policy data are taken from the dataset “EU Payments History – Regionalized and Modeled,” which contains annual data on expenditures for ERDF, CF, EAFRD, and ESF. It covers four programming periods, starting from 1989 to 1993 and ending with the 2014–2020 Program. Because some cohesion policy funds are spent after the end of the program date, the database covers years up to 2018. This data set provides the most comprehensive historical overview available to date on the annual payments made by the EU – in EUR at the current prices of the year in question within the various management funds at the NUTS-2 regional level. More information about the data used within this study can be seen in Table A1.

To enhance the coherence and consistency of the developed models, several transformations were applied to the data. Investment values were reported relative to regional GDP, either as the ratio of investments to GDP or based on regional population (investment per capita). Structural variables were transformed by normalizing them to the country's

average to account for country-wide effects. This normalization included the proportion of the population with higher and lower education levels, the share of the primary sector, the knowledge-intensive service sectors in the total gross value added (GVA) of the region, and indices of road and air accessibility. Additionally, a logarithmic transformation was applied to per capita and absolute variables to bring the data distribution closer to a normal distribution. This transformation was applied to variables such as per capita spending on research and development, gross fixed capital formation per capita, ERDF and ESF funds per capita, as well as the initial level of GDP per capita, air and road accessibility, population, population density, and the spread variable. Any missing values resulting from initially negative values were set to zero.

3.2 Methodology

A methodological approach focused on a panel structure was adopted to estimate the impact of ESIF on regional labour productivity. This approach accounts for both time variability and the specificities of each region, utilizing indicators for the eight NUTS 2 regions over the period from 2007 to 2020. The methodology aimed to capture the extensive range of variables influencing regional economic growth, referred to as explanatory variables. Additionally, the analysis included spatial spillover effects, which consider the reciprocal impacts between regions. By incorporating these elements, the approach provided a comprehensive framework to assess the influence of ESIF on regional development. An additional variable was created to quantify the spillover effects based on the economic growth of neighbouring regions. The following steps were taken: (1) the GDP growth was calculated for the investigated period for each region, (2) the differences in GDP growth between each pair of regions were determined, and (3) a weighted sum of these differences was computed for each region, using the road distances between the regions as weights. As a result, we obtained a measure for each year that was included in the model for every region to assess the significance of spillover effects. The research methodology is structured into three main stages, as shown in Figure 1.

3.2.1 Identification of determinants

The first stage uses the Least Absolute Shrinkage and Selection Operator (LASSO) to identify the main determinants of labour productivity at the regional level, where labour productivity was measured by gross domestic product relative to the number of employed people. This

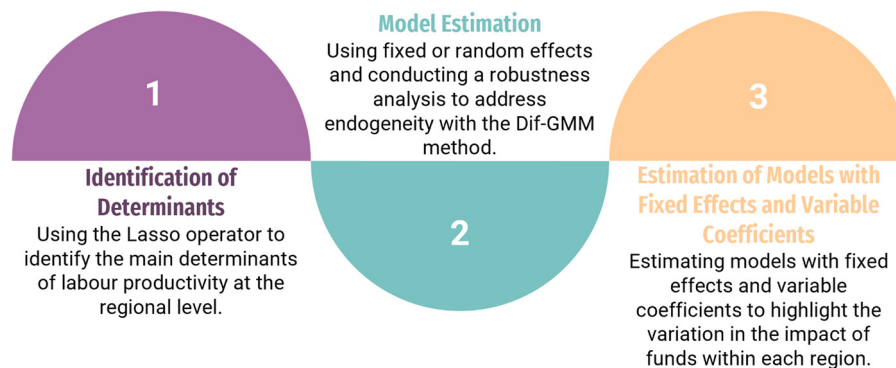


Figure 1. Research methodology.

Source: Author's own creation.

method is advantageous for variable selection and regularization, leading to more interpretable and efficient models, especially beneficial for datasets with a large dimensionality, which helps reduce the risk of overfitting.

3.2.2 Model estimation

The second stage involves estimating models based on the previously identified determinants using fixed or random effects and conducting a robustness analysis to address endogeneity with the Dif-GMM method. Addressing endogeneity in panel analysis is crucial to obtaining accurate and valid estimations.

3.2.3 Estimation of models with FE and variable coefficients

The third stage involves estimating models with FE and variable coefficients to highlight the variation in the impact of funds within each region. The choice of these models is due to the fact that each region may have its specific characteristics, such as differences in economic structure, socioeconomic framework, and demographics. These differences can influence the impact of funds on labour productivity, and therefore, the coefficients of interest in the analysis may vary from one region to another. The FE model allows for the control and inclusion in the analysis of unique characteristics for each region. Thus, we can observe and quantify how various specific factors of each region influence the relationship between European funds and regional development. Variable coefficients allow these parameters to vary by region, thus reflecting the diversity and specificity of the regional context. By using these models, we can provide a more detailed and precise perspective on how European funds affect regional development and identify significant differences between regions.

3.3 LASSO

LASSO is a regression analysis method utilized for variable selection and regularization, which improves prediction accuracy and interpretability by minimizing the sum of squared residuals (RSS) with a constraint on the coefficient estimates' absolute size (Ranstam & Cook, 2018). This method is particularly effective in reducing model complexity by setting certain coefficients to zero when the regularization parameter λ is sufficiently large, a feature absent in Ridge regression. The selection of the regularization parameter λ is crucial, as it balances shrinkage and variable selection to optimize model performance and interpretability. The Effective Bayesian Information Criterion (EBIC) is often used to determine this parameter (Lian, 2012).

3.4 Panel FE and random effects

Using a linear regression model estimated by the ordinary least squares method would fail to capture the unique data structure of panel data, neglecting the fact that specific characteristics of each region might explain some variation in the dependent variable. Acknowledging these individual effects, known as unobserved heterogeneity, is a core feature of panel models. Consequently, panel analysis decomposes the error term into two components: (i) a time-varying component and (ii) a time-invariant component within each region (Baltagi, 2010). The latter represents the influence of all region-specific factors on the dependent variable and is typically immeasurable directly. Treatment of these effects varies based on the panel model used: FE or random effects. Random effects models assume that time-invariant characteristics (individual/unobserved effects) are uncorrelated with other explanatory variables included in the model. Conversely,

FE models are employed when these individual effects are presumed to correlate with at least one of the explanatory variables.

The Hausman test compares the efficiency and consistency of the two estimators and tests the underlying assumption of the random effects model to reveal the appropriate model. The null hypothesis assumes that the random effects model is correct, meaning individual-specific effects are uncorrelated with explanatory variables, making the estimator efficient and unbiased. The alternative hypothesis suggests that FE are more appropriate, as individual-specific effects are correlated with explanatory variables, rendering the random effects estimator biased. If the null hypothesis is rejected, the FE model is preferred to account for unobserved heterogeneity and avoid omitted variable bias. If not rejected, the random effects model is chosen for its efficiency, as it captures both within and between-group variation. The Hausman test ensures model selection minimizes estimation bias while accurately reflecting variable relationships.

3.5 Difference GMM

Addressing endogeneity in panel analysis is crucial for obtaining accurate and valid estimates. Endogeneity arises when an independent variable is correlated with the error term or with other independent variables in the model, which can lead to biases in the analysis results. To address endogeneity in panel analysis, several methods can be used, including dynamic instrumental variables methods (DIF-GMM). These methods are designed to handle endogeneity in models with lagged endogenous variables or those that are time-dependent. They can correct the bias caused by correlations between independent variables and the error term.

The Difference GMM focuses on models where the dependent variable is a function of its past values and other explanatory variables (Arellano & Bond, 1991). A typical model can be expressed as:

$$y_{i,t} = \alpha y_{i,t-1} + \beta x_{i,t} + \varepsilon_{i,t}, \quad (1)$$

where $y_{i,t}$ represents the dependent variable, $y_{i,t-1}$ is the lagged dependent variable, $x_{i,t}$ includes the explanatory variables, and α and β are the coefficients to be estimated, and $\varepsilon_{i,t}$ is the error term.

To eliminate the unobserved individual-specific effects, the model undergoes a transformation where the differences are taken:

Variable	LASSO	Post-estimation OLS
log (EAFRDEMFFpc)	10.0348	10.3199**
log (ERDFpc)	2.9831	3.0211**
log (GFCF)	29.2896	29.4165**
log (GERD)	4.9112	4.9793**
EQI	-15.7593	-16.1696**
EduLow	-0.5370	-0.5364**
PrimGVA	-0.8517	-0.8623**
CONS	0.4388006	1.846910**

Table 1. LASSO's results regarding the most important determinants.

Source: Authors' own research.

Note: *, **, and *** represent 10, 5, and 1% significance level.

$$\Delta y_{i,t} = \alpha \Delta y_{i,t-1} + \beta \Delta x_{i,t} + \varepsilon \Delta_{i,t}. \quad (2)$$

4. Empirical results

Given the serious limitations of the model, especially regarding the data used (8 regions over 14 years, totalling 112 observations in a panel structure) and considering the high number of potential predictors identified (17 variables), the necessity of applying a machine learning technique becomes evident to identify variables of high importance for target variable, represented by labour productivity.

Employing LASSOPACK from STATA for the panel data, LASSO identifies significantly relevant variables for further analysis. Based on the value of the parameter $\lambda = 12.39$ (Table A2), selected based on the EBIC indicator, the selected determinants are presented in Table 1.

Thus, in Table 1, the reported values represent the estimated coefficients obtained from the LASSO regression and the subsequent post-estimation OLS model. The post-estimation OLS model was then used to obtain unbiased coefficient estimates for the selected variables. Regarding statistical significance, we acknowledge that LASSO does not inherently provide significance tests. However, for the post-estimation OLS results, all reported coefficients are statistically significant at the 5% level. According to the LASSO method, from the perspective of the funds, the determining factors were the EAFRD and the European Maritime and Fisheries Fund per capita (EAFRDEMFFpc), as well as the European Regional Development Fund (ERDFpc). The other determining factors are the proportion of the population with a low level of education (EduLow), gross fixed capital formation (GFCF), gross domestic expenditure on research and development

(GERD), governance quality (EQI), and the share of the primary sector in the gross value added (PrimGVA). Regarding the importance of the variables GERD, in a similar vein, Cardenete *et al.* (2018) also emphasized the role of investment in infrastructure and research & development (R&D) in driving sectoral productivity growth. Their study suggested that sectoral absorptive capacity plays a critical role in how effectively funds translate into productivity gains, a result consistent with our finding that GFCF significantly impacts productivity, especially in well-structured regions. The variables thus selected to explain the variations in labour productivity were used in various combinations to estimate panel data models. Although the LASSO method has significantly reduced the total number of explanatory variables, the dimensionality remains high compared to the small number of units for which the models are estimated (eight regions). The selection of the final models took into account their performance, which was quantified through significance tests and the total explained variation.

The Hausman test was used to select the type of estimated model, allowing us to choose between specifying a model with FE and one with random effects. The result of the test application imposed the selection of the alternative hypothesis (Table A3), according to which a panel model with FE is suitable. Thus, the estimation of the impact of these determinants on the variation of labour productivity was carried out using FE models but with robust standard errors, keeping only those variables in the model that have shown a significant impact on the variation of productivity:

$$LP_{i,t} = \beta_0 + \beta_1 \log(EAFRDEMFFpc_{i,t}) + \beta_2 \log(ERDFpc_{i,t}) + \beta_3 \log(GFCF_{i,t}) + \beta_4 EduLow_{i,t} + \beta_5 PrimGVA_{i,t} + u_i + \varepsilon_{i,t}, \quad (3)$$

where $LP_{i,t}$ is the dependent variable representing the labour productivity in region i at time t , $\log(EAFRDEMFFpc_{i,t})$ is the natural logarithm of the sum of EAFRD and European Maritime and Fisheries Fund per capita for region i at time t , $\log(ERDFpc_{i,t})$ is the natural logarithm of the European Regional Development Fund per capita for region i at time t , $\log(GFCF_{i,t})$ is the natural logarithm of GFCF for region i at time t , $EduLow_{i,t}$ represents the proportion of the population with a low level of education in region i at time t , $PrimGVA_{i,t}$ represents the contribution of the primary sector to the GVA in region i at time t , u_i is the unobserved region-specific effect, and $\varepsilon_{i,t}$ is the idiosyncratic error term

The empirical results, presented in Table 2, have highlighted the statistical significance of the funds from the

Variable	Coefficients
$\log(EAFRDEMFFpc)$	7.3573***
$\log(ERDFpc)$	3.2260***
$\log(GFCF)$	9.5602**
EduLow	-1.0459***
PrimGVA	1.4592*
CONS	49.6282**
<i>N</i> of observations	104
<i>N</i> of groups	8
<i>F</i> (5,91)	25.68
Prob > <i>F</i>	0.0000
<i>F</i> (FE are jointly zero)	38.71
Prob > <i>F</i>	0.0000

Table 2. Estimation of the impact of ESIF on labour productivity. Source: Authors' own research.

Note: *, **, and *** represent 10, 5, and 1% significance level.

EAFRD, the EMFF, as well as the ERDF on the variation of productivity along with the share of the population with a low level of education, gross capital formation, and the share of the primary sector in the total GVA. The results suggest that, on average, a 1% increase in the annual cumulative value of the EAFRD and EMFF funds would lead to an increase of 0.0735 percentage points in productivity at the regional level, *ceteris paribus*. In contrast, the impact of the ERDF fund on productivity is smaller, with a 1% increase in the annual value of the fund leading to an increase of 0.032 percentage points in productivity, *ceteris paribus*. This aligns with the study by Francesco and Pupo (2009), which found that EU cohesion funds contributed to reducing regional disparities but had a less pronounced effect on labour productivity, particularly in areas with weaker institutional frameworks. The agricultural sector is a major source of employment in many rural regions, which could explain the higher impact of EAFRD compared to ERDF. These findings suggest that EAFRD and EMFF investments, which directly support rural and coastal areas, have a stronger impact on productivity than the broader-focused ERDF.

Additionally, gross capital formation positively affects productivity, highlighting the importance of investment in infrastructure and technology. In contrast, a higher share of the population with low educational attainment is linked to a decrease in productivity, emphasizing the need for educational improvements to boost regional economic performance. The primary sector's share in GVA also positively influences productivity, which may reflect the foundational role of agriculture in many rural regions. The resulting equation for this final model is as follows:

Variable	Coefficients
LP(-1)	0.5863***
log (EAFRDEMFFpc)	3.1287***
log (GFCF)	7.6210**
EduLow	-0.0400**
PrimGVA	1.1370**
CONS	13.3223*
N of observations	96
N of groups	8
Wald χ^2 (5)	5725.69
Prob > χ^2	0.0000

Table 3. Empirical results of dynamic Dif-GMM estimation.

Source: Authors' own research.

Note: *, **, and *** represent 10, 5, and 1% significance level.

$$\begin{aligned}
 LP_{i,t} = & 49.6282 + 7.3573 \log(EAFRDEMFFpc_{i,t}) \\
 & + 3.2260 \log(ERDFpc_{i,t}) + 9.5602 \log(GFCF_{i,t}) \quad (4) \\
 & - 1.0459 \text{ EduLow}_{i,t} + 1.4592 \text{ PrimGVA}_{i,t} + u_i + \varepsilon_{i,t}.
 \end{aligned}$$

The robustness analysis (Table 3) aimed at addressing endogeneity and verifying the preservation of the impact of the determinants, even when changing the estimation method, highlighted the statistical significance of the two funds (EAFRD and EMFF) as well as the proportion of the population with low education, the gross capital formation, and the share of the primary sector in GVA. Furthermore, it was observed that the lag coefficient of labour productivity is statistically significant, meaning that changes or levels of productivity in the past significantly impact current productivity. In other words, the historical performance of productivity influences the present situation of productivity. This effect is positive, with past increases in productivity leading to higher productivity at present:

$$\begin{aligned}
 LP_{i,t} = & 13.3223 + 0.5863 \text{ LP}_{i,t-1} \quad (5) \\
 & + 3.1287 \log(EAFRDEMFFpc_{i,t}) + 7.6210 \log(GFCF_{i,t}) \\
 & - 0.0400 \text{ EduLow}_{i,t} + 1.1370 \text{ PrimGVA}_{i,t} + \omega_i + \varepsilon_{i,t}.
 \end{aligned}$$

The estimated models underscore a robust link between funding from the EAFRD and regional labour productivity (Table 4). This connection suggests the impact of ERDF funds in enhancing productivity at the regional level, but the results also show that this is highly variable between Romania's regions. Over the analysed period, ESIF instruments dedicated to regional development support are associated with productivity gains, yet the efficiency of these investments is notably region-specific.

In particular, the Center (RO12) and South-East (RO22) regions stand out, where the impact of EAFRD funding is especially pronounced, resulting in productivity gains exceeding 8.5 thousand lei per employee. This represents a 2.7-fold increase over the funding level. These findings indicate that specific regional characteristics, such as economic structure and institutional capacity, enhance the ability of certain areas to leverage these funds effectively, translating into meaningful productivity improvements. The economic landscape in these regions, likely coupled with a strong institutional framework, enables more effective absorption of EAFRD investments.

Conversely, the North-East (RO21) region demonstrates reduced efficiency in converting EAFRD funding into productivity growth, suggesting that regional economic and institutional factors may limit the impact of these funds. In RO21, obstacles such as structural economic restrictions or the lower quality of institutions may reduce productivity gains usually associated with this sort of funding, and thus, there is a need for tailored policy solutions that take into account regional restrictions.

In the Central Region (RO12) and the Southeast (RO22) area, results also show that GFCF investment did not have a significant contribution to labour productivity between 2007 and 2020. This suggests that, despite the financial investments in fixed assets, these regions face specific structural or institutional challenges that inhibit the full realization of productivity gains from capital investments. Factors such as labour market characteristics, industry composition, or institutional quality appear to diminish the productivity impact of GFCF in these areas, emphasizing the importance of aligning investment strategies with regional economic realities to optimize outcomes.

Similarly, Francesco and Pupo (2009) reported that southern Italian regions benefited more from EU funds, whereas northern regions showed less productivity growth, likely due to institutional inefficiencies and structural economic limitations. This suggests that the impact of EU funds on labour productivity is highly region-specific, reinforcing the need for tailored policies that address regional disparities.

In conclusion, the estimated average impact of ESIF investments on labour productivity is significant. However, it is important to observe that the impact is not uniform across all regions and that there are significant differences in the effectiveness of these investments between regions. This differentiation can be explained by the specific structural characteristics of each region, such as the population's level of education, economic composition, and degree of accessibility to resources. Thus, regions with a more diversified economic structure, a more

Region	Intercept	log (EAFRDpc)	EduLow	log (GFCF)	PrimGVA	LP(-1)
RO11	-9.614	3.826	-0.215	12.690	0.484	0.657
RO12	-15.345	9.533	0.337	-6.408	1.164	0.644
RO21	-18.865	-0.723	0.337	-6.408	1.164	0.826
RO22	-20.162	8.677	0.619	-5.458	0.212	0.660
RO31	55.856	2.299	0.067	1.976	0.009	0.272
RO32	-26.214	2.647	0.578	33.340	4.731	0.386
RO41	77.121	5.999	-1.500	11.051	-2.604	0.446
RO42	45.884	2.392	-0.925	6.661	3.807	0.143

Table 4. Estimated variable coefficients for modelling labour productivity.

Source: Authors' own research.

educated workforce, and a more developed infrastructure can better take advantage of the investments from the ESIF, generating productivity. In contrast, regions with structural or institutional constraints may experience lesser impacts. Our estimation results indicate that GDP spillover effects from neighbouring regions were not a significant factor in analysing regional labour productivity.

5. Conclusions and discussions

This article has performed a comprehensive assessment of the impact of the ESIF on the economic performance specifically labour productivity, across the eight Romanian NUTS 2 regions during the timeframe between 2007 up to 2020. The central objective was to elucidate the regional-specific impacts of these funds on labour productivity, an underexplored area in the broader discourse on EU cohesion policies. By employing a combination of machine learning and econometric methodologies, the study aimed to map out the determinants of economic performance at the regional level and explore how different variables, such as educational attainment, fixed capital formation, and spatial spillover effects, contribute to labour productivity outcomes.

The research questions have been successfully tackled since it has been established that ESIF investments contribute positively towards improvements in labour productivity although the effects are quite uneven among regions. Such discrepancies were in strong relationship with the levels of regional development and the institutional capacity. It was then concluded that the impact of ESIF effectiveness is determined by such underlying factors. Key determinants of economic performance identified included educational levels, the extent of fixed capital

investments, and the quality of local governance. These factors were consistently associated with better utilization of ESIF, emphasizing the significance of human capital and institutional quality in achieving desired economic objectives. Moreover, the ESIF effectiveness was shaped by the socioeconomic structure, technology development, and the institutional environment. Regions lacking in these aspects showed limited gains, underscoring the need for tailored approaches in policy implementation.

The findings demonstrated that ESIF had significantly boosted labour productivity, particularly in regions with better initial conditions regarding infrastructure and institutional quality. For example, regions that began with higher levels of educational attainment and more robust economic structures were able to utilize ESIF more effectively, translating these funds into substantial productivity gains. However, the analysis also revealed that the benefits of these interventions were not uniformly distributed. Some regions underwent profound transformations, while others saw only marginal improvements, highlighting the critical role of tailored regional strategies in effectively leveraging ESIF.

Overall, our results align with previous studies, reinforcing the idea that EU structural funds can enhance labour productivity, but their impact also depends on governance quality, sectoral absorption capacity, and regional characteristics. The higher effectiveness of EAFRD and EMFF compared to ERDF in our findings suggests that funds targeted towards specific sectors (e.g. agriculture, fisheries) may yield stronger productivity gains than broader economic development programs. However, as highlighted by Francesco and Pupo (2009) and Cardenete *et al.* (2018), without institutional reforms and strategic investment in human capital, the long-term benefits of these funds may remain limited.

The credibility of this study is upheld by the robust analytical methods employed, which mitigate common econometric issues such as endogeneity and data overfitting. Nonetheless, the study is not without limitations. The reliance on available data constrains the observations to 112 data points across eight regions, which might affect the statistical power of the findings. Additionally, the dynamic and evolving nature of regional policies and economic conditions over the study period introduces complexity in interpreting the impacts exclusively attributed to ESIF. An important limitation of this study is the scarcity of literature analysing the impact of EU funds on labour productivity, which results in a lack of a benchmark for comparison. This makes it challenging to contextualize our findings against a broader body of research, as most existing studies focus on GDP growth rather than productivity improvements.

Policymakers should consider region-specific strategies that address unique local challenges and opportunities. Future research could expand upon this study by exploring longitudinal impacts beyond the current timeframe and incorporating emerging economic trends and policy shifts. Additionally, investigating other qualitative aspects of development, such as social cohesion and quality of life, could provide a more holistic view of ESIF's impact.

In conclusion, this study significantly contributes to understanding how European cohesion policy can be optimized to foster equitable and effective regional development, enhancing labour productivity and overall economic resilience. By highlighting the critical role of regional characteristics, this research offers valuable insights for refining the strategic deployment of ESIF to maximize its developmental impact.

Acknowledgments

This study was carried out within the Evaluation Plan for the Partnership Agreement 2014–2020, contract no. 83996/26.08.2021, essential in assessing the contribution to economic, social, and territorial cohesion as well as the project PN 22100103.

Funding information

This research received no external funding.

Author contributions

Conceptualization, A.A.D. and M.M.M.M.; methodology, A.A.D., M.M.M.M., M.D.A. and M.B.B.; software, A.A.D.

and M.M.M.M.; validation, A.A.D., M.M.M.M., M.D.A. and M.B.B.; formal analysis, A.A.D. and M.D.A.; investigation, M.M.M.M. and M.B.B.; resources, A.A.D. and M.M.M.M.; data curation, A.A.D. and M.M.M.M.; writing—original draft preparation, A.A.D., M.M.M.M., M.D.A. and M.B.B.; writing—review and editing, A.A.D., M.M.M.M., M.D.A. and M.B.B.; visualization, M.D.A. and M.B.B.; supervision, A.A.D. and M.B.B.. All authors have read and agreed to the published version of the manuscript.

Conflict of interest statement

The authors state no conflict of interest.

Data availability statement

The data used within this study is available upon reasonable request from the authors.

References

- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297.
- Bachtler, J., & Gorzelak, G. (2007). Reforming EU cohesion policy: A reappraisal of the performance of the structural funds. *Policy Studies*, 28(4), 309–326.
- Bachtrögl, J., Fratesi, U., & Perucca, G. (2020). The influence of the local context on the implementation and impact of EU cohesion policy. *Regional Studies*, 54(1), 21–34.
- Baltagi, B. H. (2010). Fixed effects and random effects. In S. N. Durlauf & L. E. Blume (Eds.), *Microeconometrics. The New Palgrave economics collection*. Palgrave Macmillan. doi: 10.1057/9780230280816_10.
- Barro, R., & Sala-i-Martin, X. (2004). *Economic growth* (2nd ed.). MIT Press.
- Becker, S. O., Egger, P. H., & von Ehrlich, M. (2018). Effects of EU regional policy: 1989–2013. *Regional Science and Urban Economics*, 69, 143–152.
- Brandsma, A., Kancs, D. A., Monfort, P., & Rillaers, A. (2015). RHOMOLO: A dynamic spatial general equilibrium model for assessing the impact of cohesion policy. *Papers in Regional Science*, 94, S197–S221.
- Breidenbach, P., Mitze, T., & Schmidt, C. M. (2018). EU regional policy and the neighbour's curse: Analyzing

- the income convergence effects of ESIF funding in the presence of spatial spillovers. *JCMS: Journal of Common Market Studies*, 57(2), 388–405.
- Caldas, P., Dollery, B., & Marques, R. (2018). European cohesion policy impact on development and convergence: A local empirical analysis in Portugal between 2000 and 2014. *European Planning Studies*, 26, 1081–1098. doi: 10.1080/09654313.2018.1449815.
- Cardenete, M. A., Lima, M., & Sancho, F. (2018). An assessment of the impact of EU funds through productivity boosts using CES functions. *Applied Economics Letters*, 26, 872–876. doi: 10.1080/13504851.2018.1508867.
- Clement, K. (2005). Environment and sustainable development in the EU structural funds: A review of Nordic performance. *European Environment*, 15, 294–312. doi: 10.1002/EET.393.
- Crescenzi, R., & Giua, M. (2020). One or many cohesion policies of the European Union? On the differential economic impacts of cohesion policy across member states. *Regional Studies*, 54(1), 10–20.
- Durlauf, S. N. (2009). The rise and fall of cross-country growth regressions. *History of Political Economy*, 41(Suppl_1), 315–333.
- Ederveen, S., De Groot, H. L., & Nahuys, R. (2006). Fertile soil for structural funds? A panel data analysis of the conditional effectiveness of European cohesion policy. *Kyklos*, 59(1), 17–42. doi: 10.1111/j.1467-6435.2006.00318.x.
- European Commission. (2017). *My region, my Europe, our future: Seventh report on economic, social and territorial cohesion*. Publications Office of the European Union.
- Falk, M., & Sinabell, F. (2008). *The effectiveness of Objective 1 structural funds in the EU-15: New empirical evidence from NUTS 3 regions* (No. 310). WIFO Working Papers.
- Farrell, M. (2004). Regional integration and cohesion—lessons from Spain and Ireland in the EU. *Journal of Asian Economics*, 14, 927–946. doi: 10.1016/J.ASIECO.2003.09.004.
- Fidrmuc, J., Hulényi, M., & Zajkowska, O. (2019). *The elusive quest for the holy grail of the impact of EU funds on regional growth* (No. 7989). CESifo Working Paper.
- Francesco, A., & Pupo, V. (2009). The impact of European cohesion policy: A study on Italian regions. *Rivista Italiana degli Economisti*, 3, 12.
- Győrfi, A., Molnár, T., Reszkető, P., & Váradi, B. (2016). *The impact of cohesion policy on corruption and political favouritism*. Budapest Institute for Policy Analysis.
- Lian, H. (2012). Variable selection for high-dimensional generalized varying-coefficient models. *Statistica Sinica*, 1563–1588.
- Lima, M., & Cardenete, M. (2008). The impact of European structural funds in the south of Spain. *European Planning Studies*, 16, 1445–1457. doi: 10.1080/09654310802420169.
- Marzinotto, B. (2012). *The growth effects of EU cohesion policy: A meta-analysis* (No. 2012/14). Bruegel Working Paper.
- Melecký, L. (2018). The main achievements of the EU structural funds 2007–2013 in the EU member states: Efficiency analysis of transport sector. *Equilibrium*, 13(2), 285–306.
- Midelfart-Knarvik, K. H., & Overman, H. G. (2002). Delocation and European integration: Is structural spending justified? *Economic Policy*, 17(35), 321–359.
- Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T., & Busillo, F. (2013). Measuring the effects of European regional policy on economic growth: A regression discontinuity approach. *Papers in Regional Science*, 92(1), 217–233.
- Persyn, D., Diaz-Lanchas, J., Barbero, J., Conte, A., & Salotti, S. (2020). *A new dataset of distance and time related transport costs for EU regions* (No. JRC119412). Joint Research Centre.
- Psycharis, Y., Tselios, V., & Pantazis, P. (2020). The contribution of cohesion funds and nationally funded public investment to regional growth: Evidence from Greece. *Regional Studies*, 54(1), 95–105.
- Ranstam, J., & Cook, J. (2018). LASSO regression. *British Journal of Surgery*, 105. doi: 10.1002/bjs.10895.
- Rodríguez-Pose, A., & Novak, K. (2013). Learning processes and economic returns in European cohesion policy. *Investigaciones Regionales: Journal of Regional Research*, 25, 7–26.
- Rogge, N. (2019). Regional productivity growth in the EU since 2000: Something is better than nothing. *Empirical Economics*, 56(2), 423–444.
- Varga, A. (2017). Place-based, spatially blind, or both? Challenges in estimating the impacts of modern development policies: The case of the GMR policy impact modelling approach. *International Regional Science Review*, 40(1), 12–37.
- Vukašina, M., Kersan-Škabić, I., & Orlić, E. (2022). Impact of European structural and investment funds absorption on the regional development in the EU-12 (new member states). *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 17(4), 857–880. doi: 10.24136/eq.2022.029.

Appendices

Indicator	Source
Labour productivity	Eurostat
EAFRD per capita	Directorate-General for Regional and Urban Policy, EU Payments History – Regionalized and Modeled dataset
EMFF per capita	Directorate-General for Regional and Urban Policy, EU Payments History – Regionalized and Modeled dataset
ERDF per capita	Directorate-General for Regional and Urban Policy, EU Payments History – Regionalized and Modeled dataset
GFCF	Eurostat
Gross Domestic Expenditure on R&D (GERD)	Eurostat
European Quality of Government Index (EQI)	European Quality of Government Index (EQI)
Proportion of Population with Low Education (EduLow)	Eurostat
Share of Primary Sector in Gross Value Added (PrimGVA)	Eurostat
Initial GDP per capita (2007)	Eurostat
Population size	Eurostat
Road accessibility	DG Regio Regional Competitiveness Index – 2022 edition
Air accessibility	DG Regio Regional Competitiveness Index – 2022 edition

Regions analysed: 8 NUTS2 regions in Romania

Period covered: 2007–2020

Total observations: 112 observations (8 regions × 14 years)

Table A1. Data sources.

Source: Data from Eurostat; DG REGIO (EU Payments History – Regionalized & Modeled dataset; Regional Competitiveness Index 2022); European Quality of Government Index.

Lambda	L1-Norm	EBIC	R-square	FE test		
5241.22405	0.00000	705.66556	0.0000	Statistic	<i>p</i> -value	Alternative
4775.60815	3.48932	696.13488	0.1274	2.9508	0.0078	Significant effects
2268.80077	21.77004	616.51606	0.6119	Hausman endogeneity test		
1883.59826	23.47559	607.44592	0.6599	Statistic	<i>p</i> -value	Alternative
1563.79637	24.66088	601.04308	0.6941	24.7177	0.0001	One model is inconsistent
742.92997	27.86920	582.66685	0.7549	Table A3. Testing the hypothesis on random effects in modelling labour productivity.		
561.99928	29.55684	582.16015	0.7667			
425.131884	33.10793	580.57758	0.7803			

Source: Authors' own research.

Table A2. Empirical results of the LASSO method for the labour productivity indicator.

Source: Authors' own research.