Working Paper 185

AUGUST 2020

The Effects of the EU Cohesion Policy on Regional Economic Growth:

Using Structural Equation Modelling for Impact Assessment

Ambre Maucorps, Stefan Jestl and Roman Römisch

The Vienna Institute for International Economic Studies
Wiener Institut für Internationale Wirtschaftsvergleiche

The Effects of the EU Cohesion Policy on Regional Economic Growth:

Using Structural Equation Modelling for Impact Assessment

AMBRE MAUCORPS
STEFAN JESTL
ROMAN RÖMISCH

Roman Römisch, Stefan Jestl and Ambre Maucorps are Economists at The Vienna Institute for International Economic Studies (wiiw).

Research for this paper was financed by the Anniversary Fund of the Oesterreichische Nationalbank (Project No.17796). Support provided by Oesterreichische Nationalbank for this research is gratefully acknowledged.

Abstract

This paper analyses the effects of the EU Cohesion Policy (CP) on the economic growth of 276 European NUTS-2 regions between 2008 and 2016. Using a structural equation model (SEM) consisting of both a measurement component (with two latent variables) and a structural component, we estimate the impact of CP funding on the growth of GDP per capita across EU regions. The estimation also enables us to predict changes in the growth of GDP per capita based on a scenario of CP funding reallocation between member states. Overcoming the limitations of traditional linear regression, SEM modelling proves to be a promising method for impact evaluation, also allowing for the inclusion of indirect causal paths and feedback loops to depict, for example, cross-border economic spillover effects.

Keywords: Cohesion Policy, regional economic growth, structural equation modelling

JEL classification: C38, C39, R11, R12, R58

CONTENTS

1. Introduction 9

2. Literature review 11

3. Methodology and data 14

3.1. Methodological approach 14
3.2. Data 15

4. Model specification 18

5. Model estimation 22

5.1. SEM estimation and interpretation of the results 22
5.2. Scenario-based prediction 28

6. Conclusion and discussion 31

References 33

Appendix 34

TABLES AND FIGURES

Table 1 / Variables retained in the model 17
Table 2 / Linear regression with robust standard errors, time period 2008-2016 19
Table 3 / Exploration factor analysis, principal component factoring (unrotated) 20
Table 4 / Structural equation model estimation (maximum likelihood, robust standard errors) 23

Figure 1 / Initial structural equation model 21
Figure 2 / Structural equation model 22
Figure 3 / Reliance of EU regions on CP funding 26
Figure 4 / Structural laggardness of EU regions 27
Figure 5 / SEM-based simulation results 29
Figure 6 / OLS-based simulation results 30

1. Introduction

Set out in Article 174 of the consolidated version of the Treaty on the Functioning of the European Union, the Cohesion Policy (CP) aims at reinforcing the EU's economic, social and territorial cohesion by '*reducing disparities between the levels of development of the various regions and the backwardness of the least favoured regions*'. In short, the CP primarily supports convergence within the EU at the regional level, with the aim of strengthening, in turn, European integration and the EU single market. Besides convergence, the Cohesion Policy was also designed to foster competitiveness and employment, on the one hand, and territorial cooperation, on the other. With a volume of funding representing around a third of the total EU budget, the CP has been (and still is) one of the most important EU policies, alongside the Common Agricultural Policy. Over the years, the CP has shown itself to be a major instrument in supporting EU enlargement and an essential source of investment for many EU regions and countries, especially in times of financial and budgetary crisis. For some member states, the CP represents as much as 3% of their GDP (Monfort et al., 2017).

With a view to achieving balanced territorial development within the EU, Objective 1 of the CP (convergence) focuses on the least-developed regions, such as rural areas and geographically disadvantaged regions. In both the 2000-2006 and the 2007-2013 multiannual financial frameworks (MFFs), the regions eligible for CP funding under Objective 1 were those with a GDP per capita of below 75% of the EU average. For the 2014-2020 period, a twofold eligibility rule was introduced: at the national level (Gross National Income (GNI) per capita below 90% of the EU27 average[1]) and at the regional level (GDP per capita below 75% of the EU27 average). In addition, a certain amount of CP funding was also granted to so-called 'transition regions', i.e. more developed regions that had benefited under the previous MFF. The CP is financed through three structural funds (SFs): the Cohesion Fund (CF) (for Objective 1 only), the European Regional Development Fund (ERDF) and the European Social Fund (ESF). By stimulating economic growth through targeted investment in poorer regions, the CP is expected to bring benefits not only to eligible regions (e.g. through job creation), but also, indirectly, to more developed regions (e.g. spillovers through trade). The CP further gained in importance with the introduction of the euro as a common currency in many EU member states, as currency devaluation could no longer be used to compensate for economic imbalances.

However, the effects of the CP take time to become evident, as investments are spread over several years and their outputs are then reinvested to, for instance, create additional jobs, thereby yielding further positive, albeit long-term, economic benefits. At the same time, the CP has to counterbalance the cumulative growth effects of public and private investment (outside the CP) and productive activity concentration in more economically successful and attractive regions. This phenomenon has fuelled debate on the CP mainspring, and questions have been raised about whether investment money might not be better spent in more economically dynamic regions. The structure and implementation rules of the CP itself – as well as its highly (and not universally appreciated) political dimension – render it vulnerable to criticism (Barca, 2009). Hence the utility of producing evidence of its effectiveness (or indeed its lack of effectiveness) as a valuable input to the discussions on reform. While many studies have already sought to

[1] Calculated on the basis of EU figures for the period 2008 - 2010, i.e. without Croatia.

assess the impacts of the CP, the difficulty in finding a robust method to capture the CP net impacts in a long-term timeframe has naturally resulted in a wealth of publications with contrasting findings.

Providing additional insights into the effects of the CP is, therefore, of major relevance, permitting a more balanced assessment of the CP with a new, substantiated quantitative analysis. In this study, we take a methodological approach that seeks to overcome the shortcomings of previous studies, while presenting the mechanism of regional economic growth from a different perspective – that of unobserved regional factors. In particular, this study seeks to tackle the following issues:

› the vast number of (measured and unmeasured) variables that influence economic growth at the regional level (so-called ‘explanatory variables’);

› inclusion of spatial spillovers in the analysis, i.e. reciprocal effects across regional borders both within and between countries (implying that not all explanatory variables are exogenous); and

› potential measurement errors or inconsistencies (despite harmonisation procedures) in EU-wide regional datasets that consist of data reported by various regional and national statistical offices (implying that not all explanatory variables are free of errors or ‘disturbance’).

More specifically, we use structural equation modelling (SEM), a powerful modelling and estimation method that combines both factor analysis and path analysis to represent cause-effect relationships.[2] Building on the results of linear ordinary least squares (OLS) estimation (to identify observable drivers of economic growth) and factor analysis (to identify prevalent, though unobserved factors of economic growth), we produce new estimates of the impact of CP in an advanced conditional convergence model.[3] By bringing novel inputs to the CP knowledge-base, the findings of this study should eventually feed into the debate on CP reform and inform future policy-making at different governance levels: the EU level (where the CP is designed) and the national or regional level (where it is implemented). At a time when the budget of the 2021-2027 MFF (and its allocations) has yet to be finalised and new rules for EU financial support are being examined, the approach and purpose of this study become even more relevant.

The paper is structured as follows: Section 2 reviews the literature on the impacts of the CP on EU regions; Section 3 describes our step-by-step methodological approach to constructing a structural equation model, and the underlying dataset; Section 4 presents the model specification in detail; Section 5 outlines the results of the SEM estimation and, building on these estimates, the prediction of regional economic growth for a different scenario of CP funding distribution; and Section 6 concludes the study and discusses ways to further elaborate the model and exploit the full potential of SEM for regional economics.

---

2 For an example of a SEM model with latent variables, see, for instance, Wheaton et al. (1977).

3 More specifically, a model explaining growth in GDP per capita by initial GDP level (convergence), as well as other conditioning variables (including latent variables).

2. Literature review

There is a rich literature addressing the effects of the European Structural and Investment (ESI) funds on (regional) economic growth, and the issue of convergence, in particular.[4] However, there is no categorical conclusion as to whether ESI funds (or CP funds, in particular) have a significant positive effect on economic growth – and if they do, what the magnitude and longer-term sustainability of this effect could be. Actually, research studies on CP effects have mainly highlighted the local, regional and national factors that influence CP impacts and that explain their diversity (Esposti and Bussoletti, 2008). Interestingly, most research studies share the same econometric models and/or counterfactual methodological approach of regression discontinuity design, albeit with nuances (e.g. with or without spillover effects). The literature reviewed to inform our model specification focuses on more recent publications, i.e. starting from the mid-2000s, as these tend – like this study – to rely on more recent socio-economic data. Notable research findings from the late 1990s and early 2000s are nevertheless acknowledged in the studies referenced in this section, and are reflected accordingly.

An important composite study on this topic is the ex-post evaluation of the 2007-2013 CP programmes focusing on the ERDF and the CF, published by the European Commission's Directorate-General for Regional and Urban Policy (European Commission, 2016). In particular, Work Package 14 of this ex-post evaluation deals with the impacts of the CP, estimated using various methods: the QUEST III model (WP 14a), the RHOMOLO model (WP 14b), a regression discontinuity design (WP 14c) and propensity score matching (WP 14d). Performed at the national level, the QUEST III simulations revealed a substantial increase in GDP (of up to 5.4%) in those member states that receive the largest share of CP funding. Performed at the NUTS-2 level, and integrating the spillover effects stemming from accessibility-enhancing investment in transport infrastructure, RHOMOLO simulations showed that CP produces long-term positive effects on regional GDP across the entire EU. The simulations also highlighted significant intra-country variation in CP effects, mirroring the regional variation in CP funding in Europe. The regression discontinuity design study, which covered the EU15, produced an estimated additional growth in GDP per capita of 0.7 percentage points per year in Objective 1 regions (compared to those not eligible for the convergence objective) for the period 1994-2006. Similarly, the two propensity-score-matching models produced estimates ranging from 0.3 to 1.0 percentage points. Those estimates are fairly well above those of previous research studies, in particular that of Becker et al. (2010), who also investigated the impact of Objective 1 eligibility on the growth in GDP per capita over three programming periods (1989-1993, 1994-1999 and 2000-2006) through regression discontinuity design. In their study, they found a positive, statistically significant effect of Objective 1 transfers on GDP per capita growth of around 1.6 percentage points within a single programming period, and of around 1.5 percentage points when controlling for spillover effects through spatial exclusion. That study, however, encompassed considerably more NUTS-2 regions in its sample, which covered the entire EU25 (as of 2004) for the last programming period.

[4] The ESI funds currently consist of the European Regional Development Fund (ERDF), the European Social Fund (ESF), the Cohesion Fund (CF), the European Maritime and Fisheries Fund (EMFF) and the European Agricultural Fund for Rural Development (EAFRD).

Taking into account the 'absorptive capacity' of regions – measured by population education and institutional quality – Becker et al. (2013) estimated the effects of Objective 1 transfers, using regression discontinuity design with heterogeneous treatment effects. They found a positive, statistically significant effect of Objective 1 transfers on GDP per capita growth, which increases with higher endowments of human capital and better institutions. As a matter of fact, they observed that less than a third of NUTS-2 regions receiving funding under Objective 1 had sufficient absorptive capacity to transform Objective 1 transfers into economic growth. The importance of the absorptive capacity of recipient countries and regions to tap the full growth potential of CP funds had already been stressed by Bradley (2006) as 'the organizational ability to use and implement SFs efficiently and effectively'. Since that comment was made, more countries have joined the EU, further increasing the level of disparity between regions, in terms of both economic performance and structural and absorptive capacity. More recently, Crescenzi and Giua (2018) explored the heterogeneity of CP impacts across member states by adding a spatial component to regression discontinuity design. While finding an overall positive impact of CP Objective 1 funds on growth and employment at the EU level, in terms of how the CP impacts actually manifest themselves, they discovered important country effects linked to macro-national and strategic framework conditions. In particular, national macro-economic conditions, institutional conditions and quality of governance (once again, indicators of absorptive capacity), as well as acceptance (or otherwise) of the constraints and opportunities of EU policies, are believed to significantly influence the magnitude of the CP impacts observed in a given country. In fact, these impacts are estimated to be null in some member states' regions and large in others; thus specific attention needs to be paid to country-level determinants.

Departing from the vast – though methodologically narrow – research on CP impact analysis, Crescenzi (2009) investigated the pertinence and coherence of CP fund allocation through the correlation between the territorial concentration of CP funds and the (structural) socio-economic conditions for successful regional development. While the findings of this empirical analysis are not directly relevant to our research topic, it is still of interest to know more about the structural socio-economic drivers of regional economic growth, as key components of the model to be developed. Taking stock of previous research, the author categorised them as follows: factors of educational attainment, factors of productive employment and factors of demography.

The new economic geography theory has brought a highly relevant perspective to regional economic growth modelling, focusing on the interregional (cross-border) forces of agglomeration and the resultant creation of a dichotomy between core regions and peripheral regions. This phenomenon is enabled by the accumulation (concentration) of production factors through interregional (cross-border) mobility, hence the importance of concepts such as transportation and trade when studying regional economics. The diffusion of SF effects across boundaries was analysed by Dall'Erba and Le Gallo (2008), using a spatial lag econometric model. They found that, although structural funds did not have a significant effect on regional economic growth (either in EU regions as a whole or in the 'core' and 'periphery' regions separately), an interregional growth diffusion process was taking place in the core regions, i.e. those located in north-western Europe. The central location of these regions (as opposed to the peripheral location of the others) and their interconnectedness (in particular, through transportation networks) were mentioned as potential explanations for the spatially marked diffusion process. Therefore, the notion of 'accessibility' becomes decisive in the evaluation of regional development policies.

The literature review undertaken as part of this research work has revealed many different estimates of CP effects on economic growth, in terms of sign, magnitude and statistical significance; but most of them are still positive in relation to economic growth. The variety of estimates should primarily be interpreted in the light of differences in the sample data, estimation techniques and model specification used. A meta-analysis of CP impact estimations by Dall'Erba and Fang (2017) found that the geographical scope of the impact evaluation study, the exact definition and normalisation of the SF-related independent variable, the initial year of the growth period retained as a dependent variable, the control of endogeneity, and the choice of other regressors (including the interaction term) all play a role in estimation of the CP effect coefficient – whether magnifying or diminishing this estimate. Moreover, publication status and the year in which the estimation work was performed/published were also found to affect the estimate, hinting at a learning effect within the CP research community. Yet, to the best of our knowledge, there have been few impact evaluation studies covering the most recent programming periods – especially the periods 2007-2013 and 2014-2020, when the EU was enlarged to include Bulgaria, Romania and Croatia and when it was recovering from the economic and budgetary crisis. Likewise, no study has included latent variables in the estimation model, whereas the full breadth of factors influencing regional economic growth in the EU remains unknown.

3. Methodology and data

3.1. METHODOLOGICAL APPROACH

Regional economic development – which is commonly measured as GDP (per capita) growth, employment growth or productivity growth – is deemed to be influenced by a wide range of different variables: 'intrinsic' variables (i.e. development factors pertaining to the region itself), as well as 'external' variables, in particular the development of (neighbouring) regions. The latter type of variable refers to the phenomenon of spatial spillovers, when a change in a region's investment has effects not only for the region itself, but also for the economic development of other regions. The developments in other regions (caused by the initial change in the first region's investment) feed back again into the original region, and so on. These spatial feedback loops should therefore be integrated into the evaluation of CP effects on regional economic development, in order to capture both the direct and the indirect (feedback) effects of investment. It is worth noting for such an impact analysis that not all the intrinsic and external variables are necessarily observed or measured. We have developed and estimated a structural equation model (SEM), the better to identify and define the explanatory variables (including the latent ones) and to estimate their impacts (in particular that of CP funds).

The principal innovation of this study is therefore to apply SEM modelling to estimate the impacts of CP on regional economic growth. SEM is a method dating back to the first half of the twentieth century, when Charles Spearman and Sewall Wright were its main contributors; it was further developed by economists, sociologists and clinical scientists in the second half of the twentieth century (namely Robert Hauser and Arthur Goldberger, Karl Jöreskog, James Keesling and David Wiley). The method enables the inclusion of unobserved variables (latent factors represented by observed indicators and together referred to as the *measurement component* of the SEM model) and observed variables that directly (direct cause-effect path) and indirectly (indirect cause-effect path through an intermediary variable) affect an endogenous variable (together referred to as the *structural component* of the SEM model). In essence, this method overcomes (some of) the major limitations of linear regressions, as it allows for the integration of latent variables (factors), reciprocal paths and even feedback loops between variables, as well as measurement errors in the observed variables. Importantly, the SEM method represents cause-effect relationships, but does not demonstrate any causality, and SEM models must be supported by theoretical assumptions about the model specification *a priori*.

Specifically, we have applied the following procedure to elaborate and estimate a SEM model suited to our research question:

1. Establish a list of variables assumed to explain regional economic growth (taking stock of the findings of previous research studies, cf. Section 2).
2. Compute an additional variable representing economic growth in neighbouring regions (included in the model to check for the significance of spillover effects).
3. Identify the statistical significance of the coefficients linked to the variables through linear regression analysis (i.e. SEM modelling with a structural component only).

4. Carry out exploratory factor analysis on the set of explanatory variables, to assess the relevance of dimensionality reduction (i.e. SEM modelling with a measurement component only).
5. Specify the SEM model, based on the results of the linear regression (model-relevant explanatory variables), factor analysis (latent factors represented by observed indicators – i.e. measured explanatory variables) and theory-based hypotheses of covariances (hypothesised correlations between variables).
6. Estimate the SEM model.
7. Interpret the results of the model estimation.

Assess the goodness-of-fit of the model and critically review both the model specification and the estimates obtained.

3.2. DATA

Variables

Based on the literature review detailed in Section 2 of this paper, we identify the following variables as potential explanatory variables for regional economic growth:

Investment variables:

› *Cohesion policy variables:*
- Cohesion Fund expenditure as a percentage of GDP (CF)
- European Regional Development Fund expenditure per capita (ERDF)
- European Social Fund expenditure per capita (ESF)

› *Other investment variables:*
- R&D expenditure per capita
- Gross fixed capital formation per capita (GFCF)

Socio-economic structure variables:

› *Education-related variables:*
- Country-relative share of population with lower education level
- Country-relative share of population with higher education level

› *Economic structure variables:*
- Country-relative share of the primary sector in the total gross value added (GVA)
- Country-relative share of the knowledge-intensive services sector in the total GVA

› *Multimodal accessibility variables:*
- Country-relative road accessibility
- Country-relative air accessibility

› *Urbanisation proxy:*

- Population density

Control variables:

› *Convergence effect variable:*

- Initial level of GDP per capita (in euro)

› *Institutional capacity variable:*

- Quality of governance (index)

› *Size effect variable:*

- Population size

Additional explanatory variable:

› *Spillover effect variable:*

- GDP in neighbouring regions (weighted average)

This additional variable is constructed to reflect the spatial spillover effects of economic growth. It is defined as the weighted average of GDP differences in the neighbouring EU NUTS-2 regions within a 500 km radius. Within this threshold, we apply a distance decay function to weight the GDP differences of neighbouring regions: $1/distance^2$. Accordingly, higher weights are ascribed to closer regions.

Data sources

To conduct our analysis, we draw on data from various sources. Most of the data used in our analysis stem from the regional database of Eurostat. Additional data sources include the European Quality of Government Index (EQI) in its 2013 edition, developed by the Quality of Government Institute of Gothenburg University; the Regional Competitiveness Index from the European Commission's Directorate-General for Regional and Urban Policy (DG Regio), based on data from DG Regio, Eurostat, EuroGeographics and national statistical institutes. Cohesion Policy data is taken from the dataset 'Historic EU payments – regionalised and modelled'. It has yearly expenditure data for the ERDF, Cohesion Fund, EAFRD/EAGGF and ESF. It covers four programming periods, starting with 1989-1993 and ending with the 2007-2013 programme. Because some Cohesion Policy funds are spent after the programme end date, the database covers the years until 2018.

Our sample consists of 276 NUTS-2 regions across the 28 EU member states, based on the 2016 NUTS classification (i.e. all EU NUTS-2 regions, except the five French outermost regions, for which there are significant data gaps). For some regions where data were missing, estimates were made on the basis of data at a higher granulation level (NUTS-0 or NUTS-1 level). This concerned, however, a minority of variables and regions. Also, when several years were available for a specific variable, the average value over the given years was used, to avoid outlying and unrepresentative values. This concerns all investment variables, the population and population density variables, as well as our explained variable, GDP per capita growth (in constant prices).

Data transformation

In order to improve the coherence of the model and facilitate both its convergence and interpretation of the estimates, we transformed some of the data. This consisted of:

- an arithmetical transformation – dividing investment values by regional GDP (ratio of investment to GDP) or by regional population (investment per capita) and dividing socio-economic structure variables (except population density) by the country average (using country-relative variables to neutralise country effects); and
- a statistical transformation – taking the logarithm of per capita and absolute variables, in order to bring data distribution closer to a normal distribution, and applying a standardisation-like transformation to facilitate convergence of the SEM model.

Details of the data transformation can be found in the Annex.

The table below summarises the data used in our model.

Table 1 / Variables retained in the model

Type of variable	Variable
Explained	GDP per capita growth between 2008 and 2016
Investment	R&D expenditure per capita
Investment	Gross fixed capital formation per capita (GFCF)
Investment	Cohesion Fund expenditure as a percentage of GDP (CF)
Investment	European Regional Development Fund expenditure per capita (ERDF)
Investment	European Social Fund expenditure per capita (ESF)
Socio-economic structure	Country-relative share of population with lower education level
Socio-economic structure	Country-relative share of population with higher education level
Socio-economic structure	Country-relative share of the primary sector in the total GVA
Socio-economic structure	Country-relative share of the knowledge-intensive services sector in the total GVA
Socio-economic structure	Country-relative road accessibility
Socio-economic structure	Country-relative air accessibility
Socio-economic structure	Population density (as a proxy for urbanisation)
Spillover	GDP differences in neighbouring regions (weighted average)
Control	Initial level of GDP per capita (in euro, year 2008)
Control	Quality of governance (index)
Control	Population size

Source: Eurostat, Quality of Government Institute of Gothenburg University, European Commission Directorate-General for Regional and Urban Policy (DG Regio), based on data from DG Regio, Eurostat, EuroGeographics and national statistical institutes, and the Historic EU payments database. Data transformed by authors.

4. Model specification

The importance of good model specification, backed up by sound theoretical considerations, is an essential component of structural equation modelling. Indeed, when model modifications are guided more by goodness-of-fit indices than by faithful representations of socio-economic reality, it is often problematic to obtain a well-fitting model which, at the same time, allows meaningful interpretation of the results (Kaplan, 2001). A first step in specifying our model of regional economic growth is, therefore, to identify the drivers of economic growth at the regional level. This is achieved through linear regression (to identify single observed variables with statistically significant influence on our explained variable, average growth in GDP per capita in constant prices) and factor analysis (to identify 'conceptual' unobserved factors of relevance in explaining GDP per capita growth).

Linear regression

A first linear regression estimation using the above-listed (transformed) variables resulted in heteroscedastic residuals, so that robust standard errors estimation is used. The results of the latter estimation are presented in Table 2.[5]

Looking at the variables for which the regression coefficient is statistically significant at the 10% level, the results point to a convergence effect within the EU, with a negative coefficient for the initial level of GDP per capita: the lower the GDP per capita in 2008, the higher the average growth in GDP per capita between 2008 and 2016, as lagging regions catch up with more developed regions. If we now consider the Cohesion Policy support as a whole (i.e. the CF, ERDF and ESF variables), all three coefficients are statistically significant, but only ERDF and CF-related coefficients have the expected positive sign, meaning that ERDF and CF do contribute positively to GDP per capita growth. The negative sign of the ESF funding coefficient is most probably linked to the very purpose of ESF disbursement: that is, supporting employment and job creation. Thus, regions suffering from high unemployment – and from an induced slow-paced growth – are likely to be important ESF beneficiaries, generating a negative correlation between ESF funding and GDP per capita growth (reverse causality). Likewise, it can be argued that lagging regions are characterised by higher growth rates (convergence) and at the same time lower R&D expenditure, as public investment goes on other priorities (e.g. infrastructure). The opposite sign between the R&D expenditure coefficient, on the one hand, and the ERDF and CF coefficients, on the other, prompted us to analyse further the dynamics of investments. Of note is the fact that both R&D expenditure and GFCF are negatively correlated with CF, ERDF and ESF expenditure – in line with the convergence phenomenon implied by the regression analysis – R&D expenditure and GFCF are highly, positively correlated. This concurrently gives rise to a potential multicollinearity issue. One solution would be to first use factor analysis to reduce the number of explanatory variables, and then input the reduced factors into the conditional convergence model. This is exactly the approach we have taken through structural equation modelling.

[5] These regression estimates are very stable, both in magnitude and statistical significance, as shown by the inclusion of a country dummy, capital dummy or Central and Eastern Europe (CEE) location dummy in the regression (results not shown in this paper).

Table 2 / Linear regression with robust standard errors, time period 2008-2016

Explained variable:	**Average GDP per capita growth rate in constant prices**
	OLS coefficients
R&D expenditure per capita	-0.0024* (0.0013)
GFCF per capita	0.0246*** (0.0052)
CF expenditure as a percentage of GDP	0.1091*** (0.0392)
ERDF expenditure per capita	0.0025*** (0.0009)
ESF expenditure per capita	-0.0061*** (0.0014)
Country-relative population with lower education level	-0.0076** (0.0030)
Country-relative population with higher education level	0.0053 (0.0059)
Country-relative weight of the primary sector	-0.0022* (0.0012)
Country-relative weight of the knowledge-intensive services sector	-0.0054 (0.0089)
Country-relative road accessibility	-0.0039 (0.0049)
Country-relative air accessibility	-0.0008 (0.0009)
Population density	0.0022*** (0.0007)
GDP in neighbouring regions	0.0035*** (0.0009)
Initial level of GDP per capita	-0.0347*** (0.0053)
Quality of governance	0.0056*** (0.0012)
Population size	0.0032*** 0.0011
Observations	276
Number of countries	28
R-squared	0.5758

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Source: Eurostat, Quality of Government Institute of Gothenburg University, European Commission Directorate-General for Regional and Urban Policy (DG Regio) based on data from DG Regio, Eurostat, EuroGeographics and national statistical institutes, and the Historic EU payments database. Data transformed by authors.

Exploratory factor analysis

We now pool the same exogenous variables in an exploratory factor analysis, using principal component factoring to investigate the relevance of dimensionality reduction, and the size of the variance explained thereby. The results of the factor analysis are presented in Table 3.

Table 3 / Exploration factor analysis, principal component factoring (unrotated)

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
R&D expenditure per capita	0.8417	-0.1746	-0.1416	-0.0417	0.2393
GFCF per capita	0.8225	-0.3505	-0.2742	0.0974	0.1159
CF expenditure as a percentage of GDP	-0.7081	0.3420	0.1173	-0.1115	0.3554
ESF expenditure per capita	-0.7378	0.1537	-0.3938	0.3179	0.1760
ERDF expenditure per capita	-0.7837	0.2950	-0.3055	0.0892	0.1975
Country-relative population with lower education level	-0.3044	-0.3455	0.6226	0.4887	0.1616
Country-relative population with higher education level	0.5402	0.5952	-0.4379	-0.1813	0.1293
Country-relative weight of the primary sector	-0.4317	-0.5138	-0.0937	-0.3914	0.3877
Country-relative weight of the knowledge-intensive services sector	0.5608	0.6684	-0.1641	-0.0182	0.2115
Country-relative road accessibility	0.4275	0.6125	0.1760	0.2107	0.3667
Country-relative air accessibility	0.4719	0.5453	0.2440	-0.0194	0.4200
Population density	0.5405	0.4310	0.2071	0.3775	0.3367
Population size	0.2742	0.3962	0.5304	-0.4529	0.2813
GDP differences in neighbouring regions	0.5786	-0.4218	0.2812	-0.2013	0.3677
Quality of governance	0.6179	-0.6001	-0.0675	0.0861	0.2462
Initial level of GDP per capita	0.8590	-0.2879	-0.2249	0.1992	0.0890

Note: Four factors show an Eigenvalue above 1 and are therefore retained.
Source: Eurostat, Quality of Government Institute of Gothenburg University, European Commission Directorate-General for Regional and Urban Policy (DG Regio) based on data from DG Regio, Eurostat, EuroGeographics and national statistical institutes, and the Historic EU payments database. Data transformed by authors.

We observe that the first five investment variables load highly on Factor 1 – the first two positively and the last three (i.e. the SF variables) negatively. This indicates the contrasting nature of regions with high R&D expenditure and GFCF, on the one hand, and regions that benefit from Cohesion Policy support, on the other. The control variables for spillover effects, institutional capacity and initial level of GDP load highly on Factor 1 as well.

On the other hand, the higher education level variable, weight of economic sectors variables and accessibility variables load more on Factor 2, though to a lesser extent. Lower education level and share of the primary sector have a negative factor loading in Factor 2, while higher education level, share of the knowledge-intensive sector, and road and air accessibility have a positive factor loading in that same factor, indicating the contrasting features of more rural areas, on the one hand, and more urbanised, well-connected and innovative regions, on the other.

Implications for model specification

Based on the results of the regression and factor analyses, all the explanatory variables listed in Section 3 are retained in the specification of our structural equation model (SEM). We model the five investment variables as indicators of a common 'Investment' latent factor (corresponding to Factor 1 of the factor analysis) and the education, economic sectors and accessibility variables as indicators of a second 'Structure' latent factor (corresponding to Factor 2 of the factor analysis, with the exception of the lower education level variable). Building further on the theoretical framework offered by the wider literature, population density is added to the 'Structure' latent factor as an indicator of the degree of regional urbanisation (also a feature of the socio-economic structure of the region), while our control variables are introduced separately from the latent factors.

Our structural equation model therefore contains two components:

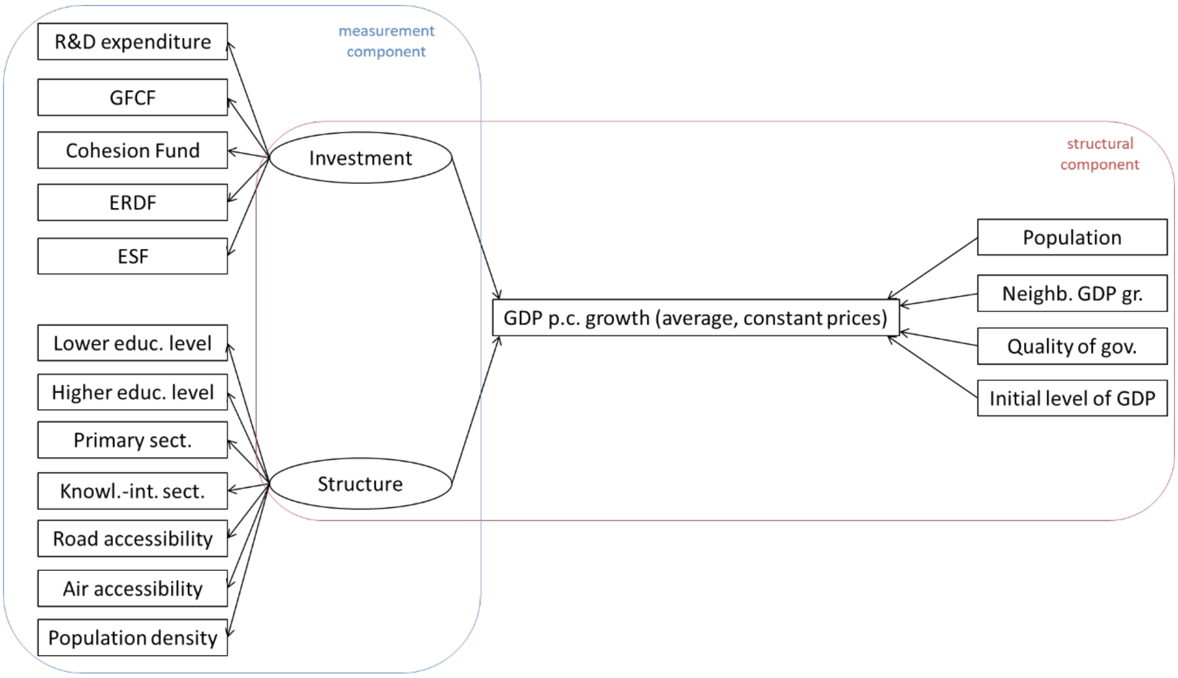
› a measurement component with two latent factors (Investment and Structure) and 12 indicators in total – this component corresponds to a confirmatory factor analysis; and

› a structural component, where GDP per capita growth is driven by the two latent factors mentioned above, as well as by population size (size effects), GDP in neighbouring regions (spillover effects), quality of governance (institutional capacity effects) and initial level of GDP per capita (convergence effects) – this component corresponds to a regression model including unobserved variables.

In the full SEM model, latent factors are interchangeably called latent explanatory variables.

Our SEM model is represented in Figure 1.

Figure 1 / Initial structural equation model



Note: rectangular forms are used for observed variables, oval forms for latent variables.
Source: authors' own elaboration.

5. Model estimation

5.1. SEM ESTIMATION AND INTERPRETATION OF THE RESULTS

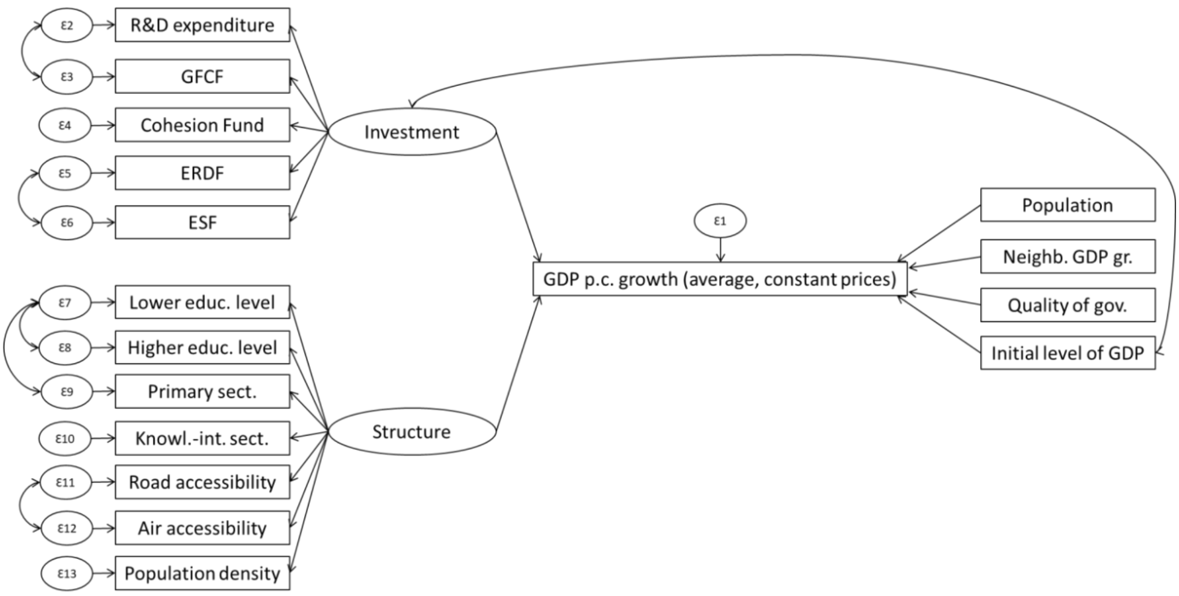
The model represented in Figure 1 above is estimated through Stata in the form described in Figure 2, i.e. with:

› error terms added to the observed endogenous variable and exogenous indicators of latent exogenous variables; and

› covariances between (error terms of) exogenous variables.

The covariance links were assumed following theoretical assumptions and preliminary correlation analysis. For instance, the share of the population with a lower education level should be (positively) correlated with the weight of the primary sector in the economy (where agricultural professions usually require less academic study); road and air accessibility should be (positively) correlated with one another (reflecting the multimodal accessibility of 'hub' regions); and the initial level of GDP should be correlated with the Investment latent variable, as a determinant of SF allocation (in particular for the Cohesion Fund).

Figure 2 / Structural equation model



Source: authors' own elaboration.

This model is estimated through maximum likelihood estimation, using robust standard errors (to correct for the heteroscedasticity of the residuals). The results are displayed in Table 4.

Table 4 / Structural equation model estimation (maximum likelihood, robust standard errors)

structural component		Coefficients
Average GDP per capita growth rate in constant prices		
	GDP growth in neighbouring regions	0.00519*** (0.00085)
	Initial level of GDP per capita	-0.02084*** (0.00246)
	Quality of governance	0.00657*** (0.00140)
	Population size	0.00299*** (0.00099)
	Investment	1 (constrained)
	Structure	-0.00715** (0.00364)
	Constant	*1.15970**** *(0.02919)*
measurement component		
R&D expenditure per capita		
	Investment	-640880.3*** (36334.9)
	Constant	5.61423*** (0.07033)
GFCF per capita		
	Investment	-350300.2*** (15292.4)
	Constant	8.46293*** (0.02975)
CF expenditure as a percentage of GDP		
	Investment	17918.8*** (1294.1)
	Constant	0.01864*** (0.00170)
ERDF expenditure per capita		
	Investment	639125.6*** (39502.3)
	Constant	3.06072*** (0.08437)
ESF expenditure per capita		
	Investment	319404.7*** (28217.5)
	Constant	2.73768*** (0.05379)
Country-relative population with lower education level		
	Structure	0.22401*** (0.04831)
	Constant	1*** (0.01326)
Country-relative population with higher education level		
	Structure	-0.42427*** (0.06007)
	Constant	1*** (0.01298)

ctd.

Table 4 / Cont.

measurement component		Coefficients
Country-relative weight of the primary sector		
	Structure	1 (constrained)
	Constant	1*** (0.04403)
Country-relative weight of the knowledge-int. services sector		
	Structure	-0.47350*** (0.06044)
	Constant	1*** (0.01275)
Country-relative road accessibility		
	Structure	-0.25458*** (0.03249)
	Constant	1*** (0.01059)
Country-relative air accessibility		
	Structure	-1.25445*** (0.19243)
	Constant	1*** (0.05363)
Population density		
	Structure	-1.90933*** (0.24043)
	Constant	5.11988*** (0.07761)
	Cov(e.R&D,e.GFCF)	0.15046*** (0.03183)
	Cov(e.ESF,e.ERDF)	0.60435*** (0.06198)
	Cov(e.PrimSec,e.EduLow)	-0.01258* (0.00694)
	Cov(e.EduLow,e.EduHigh)	-0.01485*** (0.00232)
	Cov(e.RoadAccess,e.AirAccess)	0.04116*** (0.00712)
	Cov(InitialGDP,Investment)	-4.31e-07*** (3.37e-08)
Observations		276
Number of countries		28
Coefficient of determination		0.993
Standardised root mean square residual		0.223

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
In both the structural and the measurement components, one coefficient is constrained to 1 by the model. The constrained value of the coefficient in the measurement component is used to set the metric of the latent variables.
The measurement component can be interpreted as a series of 'reverse' simple linear regressions, where each indicator is explained by the corresponding latent variable (Investment or Structure) and a constant. The order of magnitude of the coefficients is specific to the latent variable and the latent variables' respective indicators.
Source: Eurostat, Quality of Government Institute of Gothenburg University, European Commission Directorate-General for Regional and Urban Policy (DG Regio) based on data from DG Regio, Eurostat, EuroGeographics and national statistical institutes, and the Historic EU payments database. Data transformed by authors.

The coefficients and covariances are all statistically significant at the 10% level, and the majority of them even at the 1% level. In particular, the coefficients of the measurement component show a considerably higher statistical significance than those of the linear regression. The covariance between the weight of the primary sector in the regional economy and the share of the population with a lower education level is negative (contra-expectation); however, it is very significantly different from zero. All other covariance values have the expected sign.

Measurement component

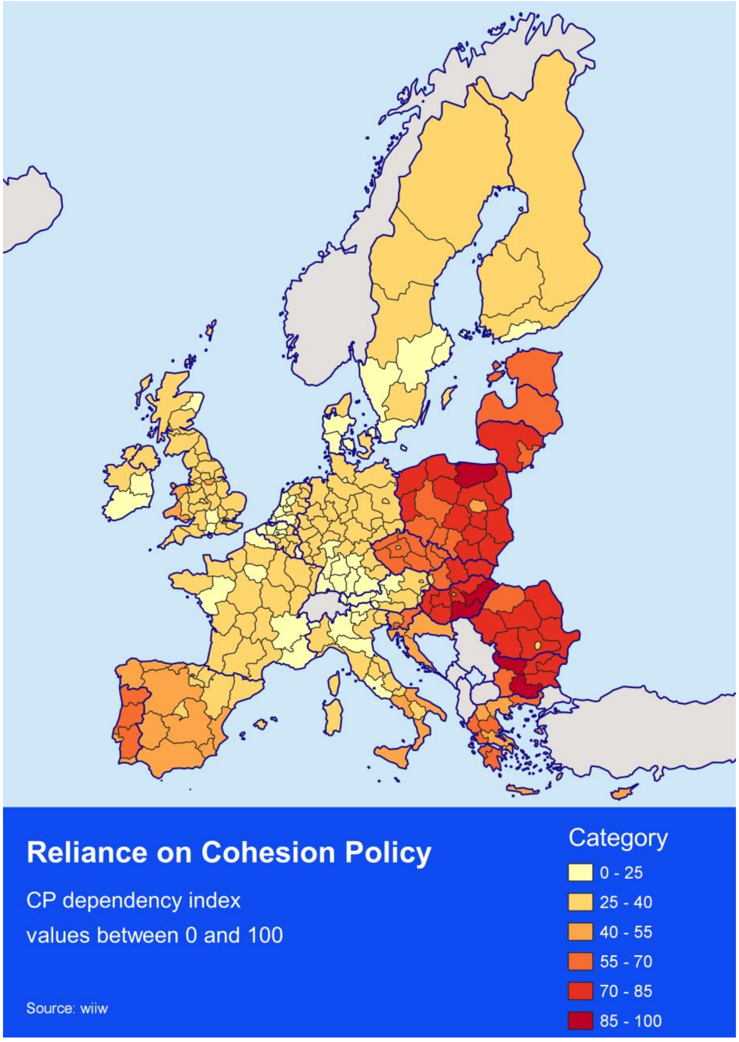
Regarding the measurement component, R&D expenditure and GFCF are negative indicators of the Investment latent variable, while CF, ERDF and ESF are positive indicators. In line with our expectations, this means that regions with a higher Investment factor are those with lower R&D expenditure and lower GFCF, on the one hand, and with higher CF, ERDF and ESF values, on the other. In that regard, the Investment factor represents the *reliance* or *dependence* of regions on CF, ERDF and ESF funding as a ‘compensation’ for lower R&D expenditure and GFCF. Lagging regions should logically demonstrate a higher Investment factor, i.e. a greater reliance on CP support.

The weight of the primary sector in the regional economy and the share of the population with a lower education level are positive indicators of the Structure latent variable, while the weight of the knowledge-intensive services sector in the regional economy, the share of population with a higher education level, road and air accessibility, and population density are all negative indicators. In line with our expectations, this means that regions with a higher Structure factor are those with a more farming-oriented economic structure and a less-educated population, as opposed to regions with an economic structure more oriented to knowledge-intensive services, a more-educated population, higher population density, and better road and air accessibility (typically urban regions). In that regard, the Structure latent variable represents the *laggardness* of the regional socio-economic structure. Lagging regions should logically show a higher Structure factor, i.e. a higher degree of structural laggardness.

The level of reliance on CP funds and the level of structural laggardness in EU regions is illustrated in the two figures below. The values of the corresponding latent variables – calculated by Stata as the means of the latent variables conditional on the observed indicators used in the model – have been rescaled to a 0-100 index, in order to improve readability.

Looking at the overall picture, we observe that the colour shades broadly match country boundaries, implying that reliance on CP is essentially a country phenomenon. The regions that rely most on the CP, as defined by our latent factor, predominantly correspond to the countries of Central and Eastern Europe (CEE), Portugal and Greece. By contrast, regions of Scandinavia, Ireland, France, the Benelux countries and the Alpine Space are least dependent on CP funding.

Figure 3 / Reliance of EU regions on CP funding



Source: authors' own elaboration.

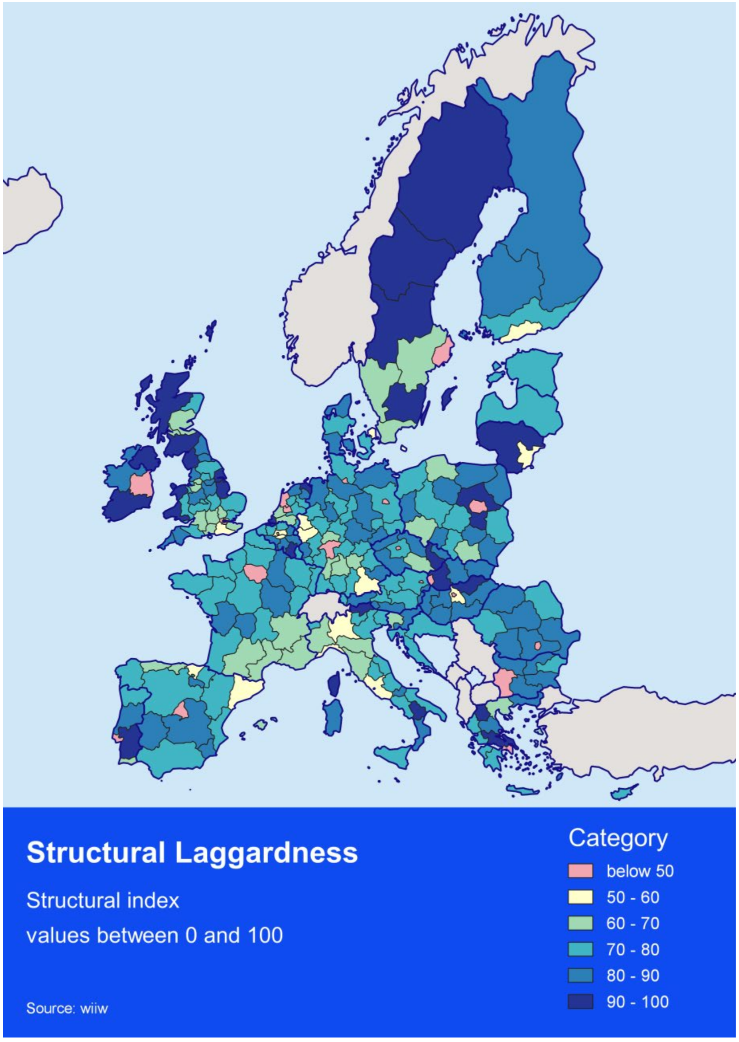
Unlike the previous figure, this figure exhibits very contrasting patterns both between and within countries, where capital regions show the smallest structural weaknesses, as defined by our latent factor, while the other regions show varying degrees of structural laggardness. Most strikingly, the traditional East-West divide is much more blurred on this figure (than, for example, on the previous one depicting CP dependence). Regions on the periphery of the EU – the southern Mediterranean regions (including island regions), Irish regions and regions in the northern UK (including island territories), northern Scandinavian regions and a few regions towards the EU's eastern external border – have particularly high scores on the structural laggardness index. These observations are supported by the new economic geography theory,[6] with two layers of core-periphery patterns:

› at the EU level, a core area of regions corresponding to the so-called 'blue banana', stretching from London to northern Italy and southern France, with the remainder of the EU being gradually 'peripheral' and 'structurally lagging'; and

[6] For details of the core-periphery model central to the new economic geography, see in particular Krugman (1991) and Fujita et al. (1999).

› at the national level, a core capital region and a wider periphery comprising by far the greatest part of the national territory.

Figure 4 / Structural laggardness of EU regions



Source: authors' own elaboration.

Structural component

Regarding the structural component, the latent 'reliance on CP funds' (Investment) variable contributes positively to GDP growth, in line with the convergence objective of the Cohesion Policy. On the other hand, the latent 'laggardness of socio-economic structure' (Structure) variable contributes negatively to GDP growth, reflecting the difficulty that regions with deficient structural capacity have in best exploiting Cohesion Policy funding. Besides, quality of governance, GDP growth in neighbouring regions and population have a positive effect on GDP growth, indicating that institutional capacity, spatial spillovers and demographic size are additional drivers of economic growth. These results are in line with the reviewed literature. Finally, the control variable 'initial level of GDP per capita' has a negative association with GDP growth, confirming the convergence effect observed in the linear regression analysis.

The coefficient of determination is very high (99%), but the standardized root mean squared residual value is quite poor (0.22), indicating that the model could be improved, for example by adding latent variables (without additional indicators) and/or causal paths to the model, backed by sound theoretical justification. This model is, however, deemed acceptable for interpreting the estimated coefficients and deriving predictions of GDP per capita growth.

5.2. SCENARIO-BASED PREDICTION

SEM prediction

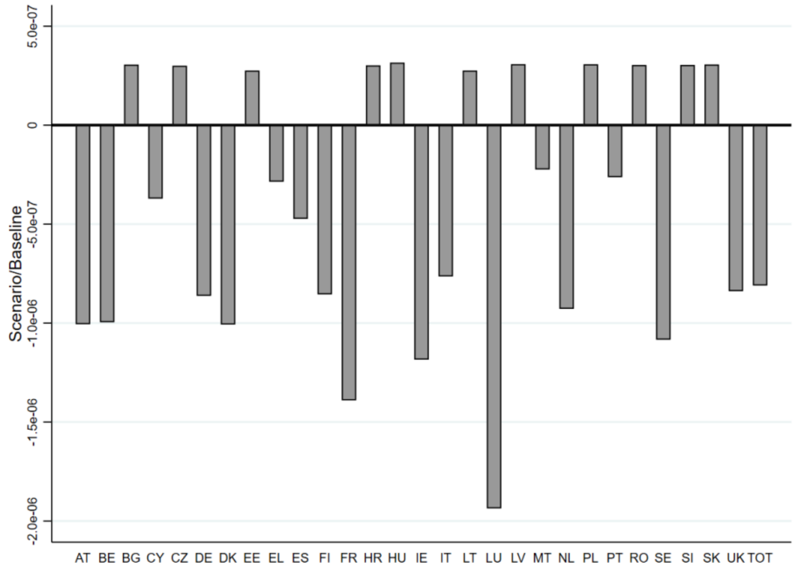
An interesting application of the SEM coefficient estimation is to predict what the GDP per capita growth would be, in different configurations of volume or allocation of CP funding. The rationale behind the prediction exercise is to examine the effectiveness of the CP, i.e. investigate whether another distribution of funding across regions would yield higher growth rates for lagging regions and speed up the convergence process. The scenario retained in this study is based on an identical volume of CP funding for as many regions as in the current MFF (in particular, UK regions are still included in the simulation).[7] However, the allocation of funding differs from the current one, to the extent that regions in CEE receive 2.5% more ERDF (total) funding (this corresponds to regions located in Bulgaria, the Czech Republic, Estonia, Croatia, Hungary, Lithuania, Latvia, Poland, Romania, Slovenia and Slovakia). The volume of additional ERDF funding corresponding to the increased allocation to CEE regions is then deducted from the ERDF allocation to the other member states' regions in proportion to the respective regions' population size. Thus, more populous regions face a larger reduction than less populous regions, but all non-CEE regions face an equivalent per capita ERDF reduction. In short, CEE regions experience an identical ERDF per capita increase in *percentage* terms, and non-CEE regions incur an identical ERDF per capita decrease in *absolute* terms. Most importantly, the total ERDF volume of funding is maintained, and as ESF and CF volumes are not changed either, the total CP funding as defined in this study is the same in the current (baseline) situation and in the ERDF reallocation scenario.

Following the changed ERDF allocation, we calculate predicted GDP per capita growth rates for all EU regions, based on the coefficients estimated through SEM (cf. Table 4). We then compute GDP per capita levels for the baseline situation and for the reallocation scenario after eight years. Finally, GDP values are aggregated at the country level and divided by the total country population, in order to compare changes in GDP per capita levels between the simulation and baseline eight years (i.e. 2016) after the year corresponding to the initial level of GDP (i.e. 2008).

The difference between the GDP per capita in the scenario simulation and the baseline situation after eight years, using SEM estimates, is displayed in the figure below for each country.

7 Considering the proposed EU budget for the next MFF, the total budget for economic, social and territorial cohesion is expected to decrease compared to the current MFF (measured in comparable price reference values). At the same time, CP funding will be shared between fewer regions (following Brexit). The average volume of CP funding per region is therefore expected to be fairly similar to the current one. More information on the proposed 2021-2027 EU budget can be found in the Conclusions of the Special meeting of the European Council (17, 18, 19, 20 and 21 July 2020), document EUCO 10/20. Available at: https://www.consilium.europa.eu/media/45109/210720-euco-final-conclusions-en.pdf

Figure 5 / SEM-based simulation results



Source: authors' own elaboration.

Positive values indicate that the scenario-based simulated GDP per capita level is higher than in the baseline situation. All CEE countries whose regions received 2.5% more ERDF funding in the reallocation scenario logically show a (quasi-identical) positive, but very small, change in GDP per capita. On the other hand, all other member states show varying degrees of negative change, as a result of the decline in ERDF funding. A larger discrepancy between the GDP per capita levels of CEE and non-CEE countries in 2016 would boost the convergence process. Yet, the graph also shows that non-CEE countries lose far more in terms of GDP per capita than CEE countries gain in the reallocation scenario. As a whole, the EU (labelled 'TOT' for 'total' in the figure) would lose out in terms of GDP per capita, as the positive changes in the CEE countries do not fully offset the negative changes in the non-CEE countries. This is because the ERDF per capita structure has changed across regions, although the total volume of ERDF funding and the total population at the EU level have not been modified.

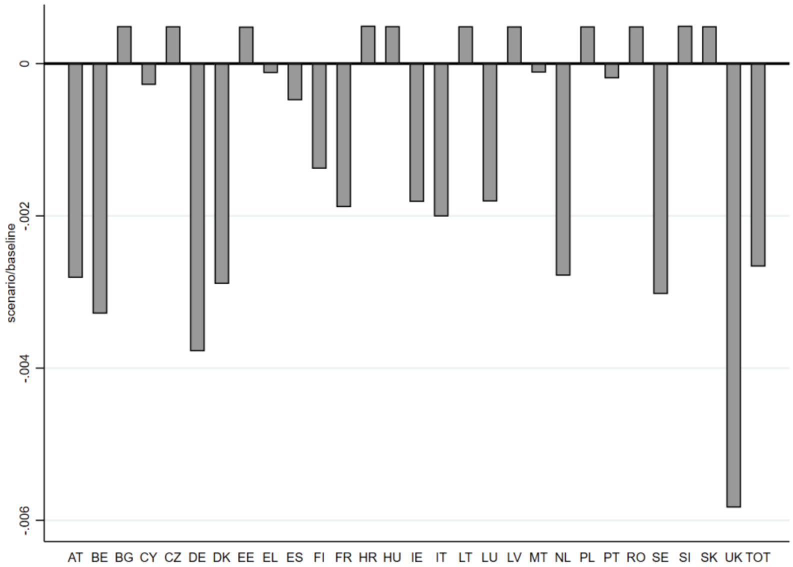
OLS prediction

The results of the scenario-based prediction using the SEM estimates are then compared with the simulation, according to the same scenario, using the linear OLS estimates displayed in Table 1. This allows us to provide a robustness check of the simulation results based on the SEM estimates.

The difference between GDP per capita in the scenario simulation and in the baseline situation after eight years, using OLS estimates, are displayed in the figure below.

These results are quite different from those of the SEM simulation, as the member states incurring the greatest 'loss' are not exactly the same as those in the SEM simulation. As in the SEM prediction, the EU as a whole experiences a lower GDP per capita level in the scenario simulation than in the baseline situation. Most notably, the magnitude of change is, in the case of the OLS simulation, significantly greater than in the SEM simulation (around 1,000 times greater). This suggests that OLS models overestimate the impact of CP funding on regional economic growth. This could be due to the omission of additional (unobserved) explanatory variables in the OLS model, which are, conversely, 'captured' by the latent variables in the SEM model. In addition, SEM estimates are deemed to be more precise, as they are not affected by multicollinearity.

Figure 6 / OLS-based simulation results



Source: authors' own elaboration.

6. Conclusion and discussion

This study proposes an innovative approach to estimating the effects of the EU Cohesion Policy (CP) on regional economic growth, using structural equation modelling (SEM). Building on the results of previous research studies that address CP impacts at the national and the regional level, our SEM model is stepwise developed: first, explanatory variables of significance are singled out through traditional linear regression; second, explanatory variables are grouped as observed indicators of latent variables through factor analysis; third, a fully-fledged recursive path model encompassing both a measurement component (consisting of two latent variables and 12 observed indicators) and a structural component (consisting of causal paths from the latent variables, control variables and spillover variable to our explained endogenous variable: average growth of GDP per capita) is specified and estimated through maximum likelihood.

More specifically, our SEM model comprehensively reflects the most important factors reported in the wider literature as explanatory of different economic growth trajectories across the EU: the structural capacity of regions to exploit the growth potential offered by CP funding – defined as a broad, unobserved factor manifesting itself in educational achievements, sectoral economic structure, accessibility and more generally level of urbanisation (approximated through population density) -, the level of investment received by regions - CP funding, of course, but also other sources of investment -, and spatial spillovers across borders. Finally, variables controlling for the region's size, the initial level of GDP per capita and quality of governance have also been integrated into the model. Unlike linear regression models, measurement errors in the observed endogenous variable and the exogenous indicators of latent exogenous variables, as well as covariances between (error terms of) exogenous variables, are explicitly specified and estimated, thereby overcoming common modelling limitations.

The estimation results reveal three very interesting features:

› First, the characterisation of a latent 'Structure' factor representing the *structural laggardness* of regions, fuelled by a more poorly educated population and a more farming-oriented economic structure, and conversely moderated by a more highly educated population, a more services-oriented economic structure, better-developed transportation systems and greater urbanisation; likewise, the characterisation of a second latent 'Investment' factor, representing the regions' *reliance on CP funding*, and driven by higher volumes of ERDF, ESF and CF funding as compensation for lower R&D expenditure and GFCF.

› Second, a statistically significant, negative effect of *structural laggardness* on GDP per capita growth and a statistically significant, positive effect of *reliance on CP funding* on GDP per capita growth. These findings support the existence and purpose of the CP, where EU funding is essential for the economic development of European regions in the absence of other abundant sources of funding, with a focus on alleviating structural weaknesses that impede the efficient use of investment for economic catch-up.

› Third, a somewhat inadequate allocation of CP funding, demonstrated by the mismatch between CP recipients (the *reliance on CP funding* is mostly a country-level phenomenon) and CP needs (the *structural laggardness* of regions is mostly a region-level phenomenon following a twofold core-periphery pattern: at the EU level and at the national level).

In line with previous studies, our model estimation also provides evidence of statistically significant size effects, spatial spillover effects, institutional capacity effects and convergence effects.

Besides, the simulation of GDP per capita growth in an alternative scenario of CP funds allocation (with more funding being transferred to CEE member states, while an identical overall volume of CP funds is maintained) has shown that economic convergence within the EU would, as expected, be expedited – but to a far lesser extent than linear OLS estimation predicts. This tends to indicate an overestimation of CP impacts through linear regression, potentially owing to omitted variables and error terms in the latter type of estimation. Importantly, both types of estimation indicate that the EU as a whole would experience a decline in the level of GDP per capita in the reallocation scenario. This further supports the idea that the CP would be more effective if funding were to be exclusively allocated at the regional level, given the (region's) lack of structural capacity to capitalise on EU investment.

In our study, structural equation modelling appears to be an innovative, sufficiently robust and fitted method to evaluate the effects of the EU CP on regional economic development. To the best of our knowledge, the SEM method has so far barely been used in economic geography – and then mainly to represent models of trade and transportation. Our study therefore represents a unique attempt to assess the recent impacts of the CP at the regional level across the entire EU, while accounting for the predominant latent factors of funding dependence and structural laggardness. It is worth mentioning a caveat related to the high sensitivity of our SEM model to changes in the sample of observations and the set of variables or paths used in the model, where minor shifts can undermine model convergence. Moreover, the calculation of the SMSR value as a goodness-of-fit index indicates that model specification can be improved – a challenging endeavour, considering the sensitivity of the model. This nonetheless makes SEM modelling a very promising method for evaluating CP impacts and producing more accurate estimates of coefficients in a well-specified setting.

Further research into the development of well-fitted SEM models could explore different aspects of regional economics – for instance, the inclusion of reciprocal paths and feedback loops between mutually reinforcing variables (to account for different types of spillover effects), the use of fine-tuned CP funding variables (e.g. CP funding along specific investment priorities) and alternative measures of economic development as an explained variable (e.g. employment or wellbeing-related variables). In addition, further research into the efficiency of the CP – i.e. the relationship between the amount of CP funding and the resulting GDP per capita gains – would surely shed more light on the 'ideal' architecture of the EU Cohesion Policy.

References

Barca, F. (2009) An agenda for a reformed cohesion policy. Independent Report prepared at the request of Danuta Hübner, Commissioner for Regional Policy.

Becker, S., Egger, P. and von Ehrlich, M. (2010) Going NUTS: The effect of EU Structural Funds on regional performance, *Journal of Public Economics*, 94(9-10), pp. 578-590

Becker, S., Egger, P. and von Ehrlich, M. (2013) Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects, *American Economic Journal: Economic Policy*, 5(4), pp.29-77. https://doi.org/10.1257/pol.5.4.29

Bradley, J. (2006) Evaluating the impact of European Union Cohesion Policy in less-developed countries and regions, *Regional Studies*, 40(2), pp.189-199. https://doi.org/10.1080/00343400600600512

Crescenzi, R (2009) Undermining the Principle of Concentration? European Union Regional Policy and the Socio-economic Disadvantage of European Regions, *Regional Studies*, 43(1), pp.111-133, https://doi.org/10.1080/00343400801932276

Crescenzi, R. and Giua, M. (2018) One or Many Cohesion Policies of the European Union? On the Diverging Impacts of Cohesion Policy across Member States, *SERC Discussion Paper 230*.

Dall'Erba, S. and Le Gallo, J. (2008) Regional convergence and the impact of European structural funds over 1989-1999: A spatial econometric analysis, *Papers in Regional Science*, 87(2). https://doi.org/10.1111/j.1435-5957.2008.00184.x

Dall'Erba, S. and Fang, F. (2017) Meta-analysis of the impact of European Union Structural Funds on regional growth, *Regional Studies*, 51(6), pp.822-832, https://doi.org/10.1080/00343404.2015.1100285

Esposti, R. and Bussoletti, S. (2008) The Impact of Objective 1 Funds on Regional Growth Convergence in the EU: A panel-data approach, *Regional Studies*, 42(2), pp.159-173, https://doi.org/10.1080/00343400601142753

European Commission, Directorate-General for Regional and Urban Policy (2016) Ex Post Evaluation of the ERDF and CF: Key outcomes of Cohesion Policy in 2007-2013. Work Package 14a: The impact of cohesion policy 2007-2013: model simulations with Quest III. Work Package 14b: The impact of cohesion policy 2007-2013: model simulations with RHOMOLO. Work package 14c: Regression discontinuity design. Work package 14d: Propensity score matching. Available at: https://ec.europa.eu/regional_policy/en/policy/evaluations/ec/2007-2013/#15

Fujita, M., Krugman, P. and Venables, A. (1999) *The Spatial Economy: Cities, regions and international trade*, Cambridge, MA: MIT Press.

Kaplan, D. (2001) Structural Equation Modeling, *International Encyclopedia of the Social & Behavioral Sciences*, pp. 15215-15222.

Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy*, 99(3), pp. 137-150.

Monfort, P., Piculescu, V., Rillaers, A., Stryczynski, K. and Varga, J. (2017) The Impact of Cohesion and Rural Development Policies 2007-2013: Model simulations with QUEST III. European Commission, Directorate-General for Regional and Urban Policy, Working Paper 05/2017.

Wheaton, B., Muthén, B., Alwin, D.F. and Summers, G.F. (1977) Assessing reliability and stability in panel models, *Sociological Methodology*, 8, pp. 84-136.

Appendix

Data transformation

In order to increase the coherence of the model and facilitate both its convergence and interpretation of the estimates, we transformed some of the data as follows:

- Dividing investment data by the region's population to use 'per capita' variables, in order to account for the region's size; this concerns R&D expenditure, GFCF, ERDF and ESF funding.
- Dividing CF funding by the region's GDP to use CF as a percentage of GDP, in order to truly reflect the economic importance of CF for the region.
- Dividing the structural variables by the country average, in order to account for country effects; this concerns the share of population with higher and lower education levels, the share of the primary and knowledge-intensive services sectors in the region's total GVA, and the road and air accessibility indices.
- Taking the logarithm of per capita and absolute variables, in order to bring data distribution closer to a normal distribution; this concerns R&D per capita expenditure, GFCF per capita, ERDF per capita and ESF per capita funding, as well as the initial level of GDP per capita, air accessibility, population, population density and the spillover variable. Missing values (because of pre-transformation negative values) have been set to zero.

IMPRESSUM

Herausgeber, Verleger, Eigentümer und Hersteller:
Verein „Wiener Institut für Internationale Wirtschaftsvergleiche“ (wiiw),
Wien 6, Rahlgasse 3

ZVR-Zahl: 329995655

Postanschrift: A 1060 Wien, Rahlgasse 3, Tel: [+431] 533 66 10, Telefax: [+431] 533 66 10 50
Internet Homepage: www.wiiw.ac.at

Offenlegung nach § 25 Mediengesetz: Medieninhaber (Verleger): Verein "Wiener Institut für Internationale Wirtschaftsvergleiche", A 1060 Wien, Rahlgasse 3. Vereinszweck: Analyse der wirtschaftlichen Entwicklung der zentral- und osteuropäischen Länder sowie anderer Transformationswirtschaften sowohl mittels empirischer als auch theoretischer Studien und ihre Veröffentlichung; Erbringung von Beratungsleistungen für Regierungs- und Verwaltungsstellen,
Firmen und Institutionen.

wiiw

wiiw.ac.at

https://wiiw.ac.at/p-5389.html