Title:

DO INSTITUTIONAL QUALITY AND EUROPEAN FUNDS AFFECT THE TRADITIONAL DRIVERS OF ECONOMIC DEVELOPMENT IN EUROPEAN REGIONS?

Abstract (1200-2000 words)

Since the seminal study by North (1990), the role of institutions in economic development is a contentious issue. From a theoretical perspective, it is widely acknowledged that institutional quality is strongly linked to economic development (Rodrik et al., 2004). On the one hand, institutional quality contributes towards improving the business environment (Acemoglu et al., 2005). On the other hand, institutional improvements can occur as consequence of economic growth (Barro, 1999). Similar conclusions have been reached in the empirical literature. In this sense, Glaeser et al. (2004) stressed the importance of not only paying attention to institutional quality measurement issues, but also the econometric techniques used to examine the role of institutions with country-level data. Other studies advocate using a regional perspective in order to offer a more appropriate level of analysis for understanding the role of institutional quality on economic development (Rodriguez-Pose, 2013).

Despite the recent interest in analyzing the relationship between the quality of local and regional government and regional performance, there is still relatively scarce evidence using data from the European regions. Only a limited number of articles have addressed this gap by examining the direct and indirect role played by institutional quality. Government quality not only affects economic growth, but also represents an indirect determinant through its effect on the efficiency of structural funds (Rodriguez-Pose and Garcilazo, 2015). The quality of institutions affects the economic performance for the whole of Europe and, particularly, in lagging regions (Rodriguez-Pose and Keterrer, 2020),thereby potentially conditioning the returns of physical and human capital as well as innovation on regional economic performance (Rodriguez-Pose and Ganau, 2022).

We contribute to this recent literature by identifying EU regions where economic growth mostly depends on the investments in R&D, infrastructures, and human capital. Hence these prove to be ideal candidate regions for European structural and investment funds. To achieve this objective, we propose estimating a latent-class economic growth model. This approach classifies the EU regions in groups according to the estimated probabilities of class-membership and offers estimations by groups in one stage (Orea et al., 2015). This information is in turn used á la Greene (2005) in order to obtain region-specific parameter estimates. The estimated parameter heterogeneity allows us to improve our knowledge about the existence of different investment patterns in the traditional drivers of economic growth and assess the direct and indirect effects of institutions in European regions. Compared to previous studies, an appealing feature of our empirical strategy is that institutional quality is considered as a determinant of the class-membership probabilities. Therefore, this methodology allows us to contrast the effect of institutions on the existing heterogeneity.

The empirical specification proposed for regional economic growth is based on the traditional model by Mankiw et al. (1992) which includes physical investment and human capital. The economy is characterized by a Cobb-Douglas production function for N regions along T periods with constant returns to scale as follows:

$$Y_{it} = A_{it} K_{it}^{\alpha_K} H_{it}^{\alpha_H} L_{it}^{1-\alpha_K-\alpha_H} \tag{1}$$

where Y_{it} is the production of the ith region in period t, K_{it} is physical capital, H_{it} represents human capital and L_{it} is the employment level. The parameter A_{it} is the region's total factor productivity level, which varies over time due to technological progress. The production function satisfies the neoclassical properties, and we assume $\alpha_K > 0$ and $\alpha_H > 0$, allowing the convergence equation to be solved.

Following the Mankiw et al. (1992)'s model, and using yearly growth rates, we obtain the convergence equation in per worker terms:

$$\Delta lny_{it} = \beta_0 - \beta_1 lny_{it-1} + \beta_2 lns_{it}^K + \beta_3 lns_{it}^H - \beta_4 \ln(n_{it} + g + \delta) \tag{2}$$

where $\Delta lny_{it} = lny_{it} - lny_{it-1}$, $\beta_0 = \beta_1 lnA_0/(1-\alpha)$, $\beta_1 = 1 - e^{-\lambda}$, $\beta_2 = \beta_1\alpha_K/(1-\alpha)$, $\beta_3 = \beta_1\alpha_H/(1-\alpha)$, $\beta_4 = \beta_1\alpha/(1-\alpha)$, $\alpha = \alpha_K + \alpha_H$, and λ represents the speed of convergence. s_{it}^K and s_{it}^H are the fractions of income invested in physical capital and human capital, respectively. Finally, the neoclassical growth model assumes that L_{it} grows at rate n_{it}, while stocks depreciate at constant rate δ and A grows exogenously at constant rate g. Therefore, this model allows us to contrast the existence of the convergence between regions, after controlling for the determinants of the steady state, named "conditional convergence."

In addition and following the idea behind the Schumpeterian endogenous growth model by Romer (1990), we consider the technological progress as a factor of economic growth in line with Rodriguez-Pose and Ganau (2022), who assumed that technology parameter A_{it} is a combination of technological adoption and quality of regional institutions. Although we focus on a semi-endogenous model given the critique by Jones (1995), we complement the above-mentioned convergence equation with the fraction of R&D investment (s_{it}^{RD}) and institutional quality (IQ_{it}). Consequently, all of our empirical analyses are performed on the basis of the following baseline specification:

$$\Delta lny_{it} = \beta_0 - \beta_1 lny_{it-1} + \beta_2 lns_{it}^K + \beta_3 lns_{it}^H - \beta_4 \ln(n_{it} + g + \delta) + \beta_5 lns_{it}^{RD} + \beta_6 lnIQ_{it} + \eta_i + v_{it} \tag{3}$$

where η_i is a regional-specific effect and v_{it} is the traditional identically and independently distributed disturbance term. We also extend this specification adding the length of the road network in order to examine the role of transport infrastructures as an economic growth driver, but using a smaller sample due to the lack of data in several regions.

Notice that all parameters in equation (3) are common to all regions and time-invariant, which is likely a strong assumption. We propose using a Latent Class Model (LCM) to estimate equation (3) in order to obtain heterogeneous parameters across regions (Orea et al., 2015). LCMs separate the sample into a finite number of classes (regimes) and estimate a specific economic growth function for each class in a single stage. The allocation of regions to a particular class relies on the estimated class membership probabilities that reflect the uncertainty researchers may have about the true partition of the sample.

Encompassing in θ_j all the parameters associated with class *j*, the conditional likelihood function of an observation belonging to class *j* can be denoted by $LF_{ijt}(\theta_j)$. The unconditional likelihood is then obtained as the weighted sum of the *j*-class likelihood functions, where the weights are the probabilities of class membership. That is:

$$LF_{it}(\theta) = \sum_{j=1}^{J} LF_{ijt}\left(\theta_j\right)P_{ijt}, \qquad 0 \leq P_{ijt} \leq 1, \qquad \sum_{j=1}^{J} P_{ijt} = 1 \quad (4)$$

where $\theta = (\theta_1, \ldots, \theta_j)$ and the class probabilities P_{ijt}, which must be between 0 and 1, can be parameterized as a multinomial logit model:

$$P_{ijt}\left(\delta_j\right) = \frac{\exp(\delta'_j q_{it})}{\sum_{j=1}^{J}\exp(\delta'_j q_{it})}, \qquad j = 1, \ldots, J \ , \quad \delta_J = 0 \quad (5)$$

where q_{it} is a vector of variables that are related to the latent problem being modelled, and δ_j is a vector of parameters to be estimated.

The overall likelihood function resulting from (4) and (5) is a continuous function of the vectors of parameters θ and δ, and can be written as:

$$ln\, LF\ (\theta, \delta) = \sum_{i=1}^{N}\sum_{t=1}^{T} ln\, LF_{it}\ (\theta, \delta) = \sum_{i=1}^{N}\sum_{t=1}^{T} ln\left\{\sum_{j=1}^{J} LF_{ijt}\left(\theta_j\right)P_{ijt}\ (\delta_j)\right\} \quad (6)$$

Maximizing the above likelihood function gives asymptotically efficient estimates of all the parameters. The estimated parameters can be used to compute the posterior probabilities using the following expression:

$$P(j|it) = \frac{LF_{ijt}(\hat{\theta}_j)P_{ijt}\ (\hat{\delta}_j)}{\sum_{j=1}^{J} LF_{ijt}(\hat{\theta}_j)P_{ijt}\ (\hat{\delta}_j)} \quad (7)$$

These posterior probabilities can then be used to assign each region to a particular class, i.e. each firm is allocated to the class with the highest posterior probability.

In our application, we assume that institutional quality determines the probability of pertaining to a different class, or in other words represents the source of the heterogeneity between regions. Therefore, this model allows us to analyze the direct effect of institutional quality on regional development as well as its indirect effect through investment, innovation, human capital and transport infrastructures. Estimations are carried out via the maximum likelihood method with fixed and temporal effects.

Some of the determinants of regional economic growth can be considered as endogenous. Therefore, it is necessary to estimate by applying a method of instrumental variables (Wooldridge, 2002). The most appropriate instruments are obtained through the residuals from the auxiliary regressions, in which the endogenous variables are regressed with respect to exogenous variables that meet the characteristics of a valid instrument. We assume that the most appropriate exogenous variables to explain investment, innovation, human capital and transport infrastructures are the European funds lagged three periods[1] and the growth rates of the population. We consider as an exogenous variable for institutional quality its four years lags[2] and the population growth rate. These auxiliary regressions allow us to understand the role of European Funds on regional development through their contribution to the drivers of economic growth and the existence of synergies between them.

The empirical application is performed on a sample of 230 EU NUTS-2 regions during the period 2007-2018. Regionalized data on EU funds are taken from the Historic EU payments dataset provided by the Open Data Portal for the European Structural Investment Funds (ESIF). The dataset includes regionalized EU payments for the different funds: (i) the European Regional Development Fund (ERDF), (ii) the European Social Fund (ESF), (iii) the Cohesion Fund (CF), (iv) the European Agricultural Fund for Rural Development (EAFRD), (v) the European Maritime and Fisheries Fund (EMFF), and (vi) the Youth Employment Initiative (YEI), and (iii) the Fund for European Aid to the Most Deprived (FEAD).

Data on regional gross domestic product (GDP) and gross fixed capital formation (GFCF) at constant prices, population, and employment come from the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO). Human capital – measured as the percentage of population between 25 and 64 years with secondary education –, intramural R&D expenditures, and kilometers of motorways and other roads are taken from Eurostat and the QoG EU Regional dataset (Charron et al, 2020). The quality of regional institutions is measured though the European Quality of Government Index (Charron et al, 2021). Descriptive statistics are presented in the following table.

Table: Descriptive Statistics

Variable	Mean	Std. dev.	Min	Max
Growth of GDP per capita (%)	0.01	0.03	-0.18	0.14
GDP per capita (€)	2,6043.30	13,693.56	3,646.52	94,780.27
Population growth (%)	0.00	0.01	-0.06	0.07
Investment (% of GDP)	0.21	0.04	0.07	0.69
EU Funds (Million €)	207.30	258.18	0.32	2,305.32
Institutional Quality	0.11	0.99	-2.80	2.82
Human Capital (%)	26.00	8.75	6.80	57.10
R&D (€ per capita)	480.90	551.23	3.00	3,884.30

Note: Descriptive statistics based on a sample of 2,294 observations for 230 regions.

[1] European funds must be spent by the second or third year after their allocation. This is known as the N + 2 or N + 3 rule.

[2] At the end of a legislature, it is possible that a change of government will lead to changes in the institutional quality

Our preliminary results show that the quality of government is significant in the classification of regions in the latent groups, and that its effect on the role of investment varies across regions. These results have important implications for policy-makers, as the increase in government quality could enhance the impact of European investments. Finally, several counterfactual exercises are also performed to analyze the potential effect of the European funds provided to mitigate the economic losses caused by the current pandemic of COVID-19.