Received: 29 April 2023 | Revised: 6 October 2023 | Accepted: 22 January 2024

DOI: 10.1111/grow.12711

growth and change

WILEY

ORIGINAL ARTICLE

Addressing spatial dependence when estimating technical efficiency: A spatialized data envelopment analysis of regional productive performance in the European Union

Julián Ramajo[1] | **Miguel A. Márquez**[1] | **Geoffrey J. D. Hewings**[2]

[1]Facultad de Ciencias Económicas y Empresariales, Universidad de Extremadura, Badajoz, España

[2]Regional Economics Applications Laboratory, University of Illinois, Urbana, IL, USA

Correspondence
Julián Ramajo.
Email: ramajo@unex.es

Funding information
Ministerio de Ciencia e Innovación, Grant/Award Number: PID2019-109687GB-I00

Abstract

Applying a spatialized data envelopment analysis, this paper estimates and analyzes the efficiency of European Union NUTS-2 regions during the period 2000–2014. The space-dependent efficiency scores estimated with the proposed model show a bimodal distribution that is not detected by the aspatial approach. The results confirm the crucial role of location of production units, offering important new insights on both the causes of regional disparities in labor productivity and the observed polarization of the European distribution of per capita income. The findings further suggest significant differences for the two subgroups (spatial clusters) of regional economies found in the correlation analysis between the efficiency scores and the components of the EU Regional Competitiveness Index. Consequently, policies suitable for one group of regions may not be suitable or appropriate for the other.

KEYWORDS

data envelopment analysis, European Union, production frontier, regions, spatial dependence, technical efficiency

1 | INTRODUCTION

In their introduction to a collection of articles addressing the Rebuilding Macroeconomic Theory Project, Vines and Wills (2018) noted that four main changes to the core model were recommended. These were the needs to (1) emphasize financial frictions; (2) place a limit on the operation of rational expectations; (3) devise more appropriate microfoundations and (4) include heterogenous agents. While the latter concern was addressed in terms of heterogenous consumption behavior of households of different size or income levels, none of the authors considered the role and impact of heterogenous space. No attention was directed to the problem of differences in economic structure and performance at the sub-national (regional) levels and the way in which ignoring this spatial heterogeneity might create problems for the effective performance of macroeconomic models (the derived bias could be both positive and negative, but anyway the induced deviation could be important).

This paper explores the role of the spatial dimension in the context of concerns about the increase of global income inequality (Alvaredo et al., 2018), particularly the lack of income convergence between countries (Johnson & Papageorgiou, 2020) and the rising imbalances within countries (Blanchet et al., 2019). These general trends are of special relevance for the case of the European Union (EU), where the reduction of regional disparities is explicitly included in the Treaty of the Functioning of the EU. Notwithstanding the existence of the EU Cohesion Policy as a tool to address regional disparities, the empirical evidence reveals that while inequalities *between* the European countries are decreasing, regional income disparities *within* countries have been increasing (Alcidi, 2019).

Currently, one of the most significant challenges faced by regions aiming to boost their incomes is the need to increase their labor productivity. In Europe, some regions are struggling with low productivity levels, making it harder for them to catch up and close the income gap with their more prosperous counterparts (Widuto, 2019). In recent years, researchers have identified a complex set of spatial factors that are interdependent and serve as fundamental drivers of productivity dynamics (Tsvetkova et al., 2020). Since regional well-being is closely tied to productivity, studying the sources of per capita income disparities across countries has been a focus of researchers seeking to increase aggregate output and productivity (Syverson, 2011). Despite the importance of the spatial dimension of productivity, mainstream research has often overlooked the spatially heterogeneous distribution of this variable.[1] This is a significant oversight, as regional disparities in the EU may be linked to the productivity gap in Europe (Bruno et al., 2019). Therefore, it is crucial to gain a better understanding of the role played by spatial factors in productivity dynamics, and to diagnose the determinants of the ongoing divergences in the EU. This will enable policymakers to assess the performance of regional economies in Europe and implement targeted measures to address productivity challenges and reduce regional disparities.

Our research provides a new contribution to the Data Envelopment Analysis (DEA) literature, the Spatial DEA (SpDEA) model that can assess the efficiency of spatially referenced decision-making units (DMUs). We apply our proposed SpDEA specification to 263 European NUTS-2 regions over the period 2000–2014. It was found that the results of the two models (DEA and SpDEA) are different; in fact, after controlling for geographical externalities, by construction, the average regional efficiency increased with respect to the standard DEA results. Another important finding is that the shape of the densities associated with the efficiency scores estimated by the SpDEA model for each year (2000, 2007, and 2014) shows a tendency toward a bimodal distribution; however, this twin-peakedness is not detected if the biased scores from the standard (aspatial) DEA are used.

The relevance of these results stems from the fact that they promote the need to consider this important dimension in the context of the European policy goals related to competitiveness. The competitive positions of the European regions relate to the mechanisms by which regions change their technical efficiencies; critically, these efficiencies are related to those in nearby regions. The research suggests that the spatial distribution of efficiencies could play a fundamental role in the competitiveness of the European regions, having importance for the spatial design of the European policy in terms of regional development and productivity. In addition, there is an economic rationale behind our results that is coherent with the empirical analyses carried out previously in terms of regional productivity and convergence for the case of world countries (Badunenko et al., 2013, 2018; Henderson & Russell, 2005) or the European regions (Rogge, 2019), shedding light on the causes of the observed regional performance inside the EU.

The rest of the paper is organized as follows. The methodological framework of the analysis and the proposed SpDEA approach are presented in Section 2. The data and variables are described in Section 3. Also, in this empirical part, we estimate the standard DEA and the SpDEA model to account for regional spillovers. Section 4 summarizes the results and offers some conclusions.

2 | SPACE-DEPENDENT EFFICIENCY SCORES: THE SPATIALIZED DATA ENVELOPMENT ANALYSIS

This section starts by placing the methodological contribution of this work in context showing recent developments in the nonparametric frontier literature. Thereafter, we overcome the limitation detected in the standard DEA technique when dealing with efficiencies by means of the proposed SpDEA model.

In this framework, the technical efficiency of a regional economy would be the degree to which its production level approaches the maximum output obtainable with its disposable inputs (Färe & Lovell, 1978). Hence, a measure of aggregate technical efficiency would inform about how regional economies in the EU are operating with their inputs; this would enable comparison of the characteristics of the regions operating near the frontier with those that are less efficient, facilitating the formulation of appropriate policies to address these important challenges with the reasonable expectation that there would be an important dividend in terms of economic performance at the EU-wide level.

A robust finding of the large and growing literature using regional data is that agglomeration externalities and their associated spatial spillovers shape regional performance (e.g. Davis et al., 2014; De Groot et al., 2016; Fracasso & Vittucci-Marzetti, 2018; Fritsch & Kublina, 2018; Kerr & Kominers, 2015). The study of these spatial effects on regional convergence has received theoretical and empirical attention (see, for example, Badinger et al., 2004; Cosci & Mirra, 2018; Gobillon & Magnac, 2016; Kubis & Schneider, 2016; Ramajo et al., 2008).

In recent decades, applied researchers have become increasingly interested in the analysis of efficiency across regional economies in order to determine whether or not the use of regional factor endowments is contributing to regional convergence (Maudos et al., 2000; Ramajo & Hewings, 2018; Rogge, 2019). This interest is related to the important role that changes in efficiency could have on changes in productivity (Solow, 1957) affecting the regional convergence process. One strand of this literature has focused on the analysis of regional technical efficiency (Fritsch & Slavtchev, 2011; Tsekeris & Papaioannou, 2018). Regions face different contextual variables related to the location in which they operate and some of these characteristics, that are beyond the control of local authorities, affect their aggregate economic performance.

The underlying sources of agglomeration economies constitute the main mechanisms through which efficiency within a region is influenced. These sources have been a longstanding focus in economics (for a comprehensive overview, refer to Cohen & Morrison, 2009; Puga, 2010), with their micro-foundations described by Marshall (1920) and Rosenthal and Strange (2004). Marshall (1920) outlined three mechanisms through which agglomeration operates: labor market integration, knowledge diffusion, and shared resource utilization. Additionally, natural advantages, the presence of a local market, consumption patterns, and rent-seeking behaviors are other potential and equally significant underpinnings of agglomeration externalities (Rosenthal & Strange, 2004).

When one or more of the aforementioned basic sources are in operation, the resulting agglomeration economies are not confined to a single firm; rather, they can impact all firms located in the same vicinity. As a result, the specific mechanisms driving agglomeration economies may differ from one region to another. As we shift our focus to the regional context, the advantages traditionally linked to firms are gradually being supplanted by benefits that are more appropriately considered within a broader regional perspective. Essentially, agglomeration economies are factors associated with the proximity or concentration of businesses and workers, resulting in cost savings in production (Cohen et al., 2019). These cost savings, related to the existence of intra-regional economies of scale, can impact the overall capacity utilization of a region influencing regional technical efficiency. Moreover, agglomeration economies have effects that extend beyond their immediate location and can impact neighboring regions.

Our research highlights that the impact of regional agglomeration effects can extend beyond the geographical confines of a specific region, potentially influencing neighboring areas. This perspective explains the varying efficiency levels observed among regions that may be similar in nature but situated in different neighborhoods. In essence, the crucial factor in this phenomenon is the presence of geographical externalities: agglomeration economies that are dependent on immediate spatial proximity (Parr, 2002). This means that the reach of agglomeration economies extends well beyond the traditional boundaries of a single region, as discussed by Phelps et al. (2001) and Parr (2002). The micro-sources of geographical externalities between regional economies are based on the previously mentioned micro-foundations (Marshall, 1920; Rosenthal & Strange, 2004), but incorporating the regional foundations of agglomeration economies, where the spatial interactions between regional economies is a fundamental component. Thus, below, we outline some key mechanisms of geographical externalities between regional economies that highlight how agglomeration economies can lead to increased efficiency and competitiveness for firms in neighboring regions.

First, knowledge spillovers in neighboring regions can result from the diffusion of ideas, innovations, and expertise between firms and institutions located in close proximity. In addition, proximity fosters cultural and social exchanges between regions, influencing the diffusion of knowledge and ideas (knowledge diffusion) and shaping the cultural and social. Secondly, geographical proximity allows for more efficient supply chain integration. Firms in neighboring regions can collaborate in the production process, share suppliers, and reduce transportation costs, ultimately increasing the competitiveness of both regions. Thirdly, industries often cluster in regions, and adjacent regions may witness the emergence of complementary industry clusters. Furthermore, these neighboring industries can derive advantages from shared resources, access to a skilled labor force, and knowledge exchange, ultimately bolstering their overall competitiveness. Fourthly, neighboring regions may collaborate on infrastructure projects, such as transportation networks or research facilities, which reflects the sharing of resources and infrastructure within agglomerated areas. Fifthly, proximity between regions can enhance market access for firms, enabling them to tap into each other's consumer markets more efficiently. Finally, regional policies can affect neighboring regions. Regional policymakers in neighboring regions may coordinate policies and incentives to attract firms and investments, leading to policy spillovers.

Therefore, these mechanisms of geographical externalities between regional economies emphasized that geographical proximity between regions becomes a crucial factor in shaping regional efficiency. However, it is also important to note that agglomeration effects can have either a positive or negative impact on regional technical efficiency. When examining regional efficiency, the effects stemming from these various sources of externalities can culminate in a combined aggregate effect at the regional level (Duranton & Puga, 2004). Consequently, neighboring regions may engage in competition to outperform each other in terms of economic growth and innovation. For example, competition among neighboring regions can result in rent-seeking behaviors, where regions actively seek to attract businesses and investments through competitive policies, incentives, and strategies.

In conclusion, the overall estimated agglomeration effect on regional technical efficiency could be caused for the most part by the micro-foundations of agglomeration economies, by the micro-sources of geographical externalities between regional economies or by a combination of both types of micro-sources.

Consequently, the estimation and assessment of efficiency in regional production should consider the spatial dimension of the DMUs analyzed. Spatial dependence in regional production efficiency refers to the correlation between the efficiency level of the regions and those of the neighboring regions.

To measure technical efficiency, one of the most popular nonparametric approaches is the DEA, introduced by Charnes et al. (1978, 1981). When using this technique, until recently, most existing nonparametric empirical studies did not consider the existence of spatial dependence in the estimation of the efficiency frontier. For example, in the surveys of Emrouznejad & Yang, 2018, Moradi-Motlag & Emrouznejad, 2022, or Ray, 2020, no mention is made of the spatial dimension of the data and its implications. This lack of spatial consideration could be causing biases in the estimation of the technical efficiency scores, hiding relevant information on the performance of the DMUs and providing distorted results in the benchmarking analysis of productive performance.

2.1 | The standard (output-oriented) DEA method

Following the seminal work of Färe et al. (1994), this paper addresses the issue of the relative technical efficiency by means of the DEA approach that does not require the specification of the functional form of the production function being estimated. In order to measure the efficiency of regions, we define a set of p inputs $\boldsymbol{x} = \left(x_1, x_2, \ldots, x_p\right) \in \mathbb{R}_+^p$ that are used to produce a vector of r outputs $\boldsymbol{q} = (q_1, q_2, \ldots, q_r) \in \mathbb{R}_+^r$. Then, the technology set of all feasible input-output combinations $(\boldsymbol{x},\boldsymbol{q})$ can be defined as:

$$\psi = \left\{(\boldsymbol{x}, \boldsymbol{q}) \in \mathbb{R}_+^{p+r} \mid \boldsymbol{x} \text{ can produce } \boldsymbol{q}\right\} \tag{1}$$

The (unconditional) output-oriented Farrell-Debreu technical efficiency DEA score, $\hat{\lambda}_{DEA}$, of a DMU can be obtained by solving the following linear programming problem:

$$\hat{\lambda}_{DEA}(\boldsymbol{x}, \boldsymbol{q}) = sup\{\lambda > 0 \mid (\boldsymbol{x}, \lambda\boldsymbol{q}) \in \hat{\psi}\} \tag{2}$$

where $\hat{\psi} = \left\{(\boldsymbol{x}, \boldsymbol{q}) \in \mathbb{R}_+^{p+r} \mid \boldsymbol{q} \leq \sum_{i=1}^{n} \gamma_i \boldsymbol{q}_i, \boldsymbol{x} \geq \sum_{i=1}^{n} \gamma_i \boldsymbol{x}_i, \sum_{i=1}^{n} \gamma_i = 1, \gamma_i \geq 0\right\}$ is the attainable set estimated from an observed random sample of DMUs $\{(\boldsymbol{x}_i,\boldsymbol{q}_i)|i = 1,2,\ldots,n\}$, and λ is the efficiency parameter to be evaluated for the productive unit operating at level $(\boldsymbol{x},\boldsymbol{q})$. Using this definition, $\hat{\lambda}_{DEA}(\boldsymbol{x}, \boldsymbol{q}) = 1$ denotes an efficient production unit, while $\hat{\lambda}_{DEA}(\boldsymbol{x}, \boldsymbol{q})^{-1} < 1$ implies that the corresponding DMU is inefficient.

2.2 | The influence of contextual variables and the conditional DEA estimator

To address spatial dependence inherent in geographically referenced data, our DEA specification will incorporate spatial autoregressive terms as external conditioning factors. This approach will allow us to estimate regional efficiency scores using the conditional nonparametric frontier method developed by Daraio and Simar (2005, 2007). Our spatialized DEA model provides a means to explore the potential impact of interregional spillovers on regional production processes. Un-like the parametric production approach, where spatial effects are associated with specific parameters that measure the relevance of geographical externalities across cross-sectional units, our model considers that the effects of spatial factors are not always one-directional (positive or negative). Instead, they may have a variable influence on regional productivity performance. By incorporating spatial factors in our analysis, we can better understand the complexities of regional productivity and identify opportunities to enhance efficiency and reduce disparities.

Using the conditional frontier method of Daraio and Simar (2005, 2007) - a summary explanation of this approach is provided in the Appendix-, the conditional DEA (cDEA) output measure of technical efficiency can be obtained as:

$$\hat{\lambda}_{cDEA}(\boldsymbol{x}, \boldsymbol{q}|\boldsymbol{z}) = sup\left\{\lambda > 0 \mid \hat{S}_{Q|X,Z}(\lambda \boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}) > 0\right\} \tag{3}$$

where by construction $\hat{\lambda}_{DEA}(\boldsymbol{x}, \boldsymbol{q})^{-1} \leq \hat{\lambda}_{cDEA}(\boldsymbol{x}, \boldsymbol{q}|\boldsymbol{z})^{-1} \leq 1$ since for all z, $\psi^z = \{(\boldsymbol{x},\boldsymbol{q})|\boldsymbol{Z} = z, \boldsymbol{x}$ can produce $\boldsymbol{q}\} \subseteq \psi$. The estimated survival function $\hat{S}_{Q|X,Z}$ in (3) is more difficult to evaluate than in the unconditional case (the support set $\hat{\psi}$ in Equation 2), because it requires the use of smoothing techniques for the external variables in z (see Bădin et al., 2010):

$$\hat{S}_{Q|X,Z}(\boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}) = \frac{\sum_{i=1}^{n} \mathbb{I}\left(\boldsymbol{x}_i \leq \boldsymbol{x}, \boldsymbol{q}_i \geq \boldsymbol{q}\right) K_h\left(\frac{z_i - z}{h}\right)}{\sum_{i=1}^{n} \mathbb{I}(\boldsymbol{x}_i \leq \boldsymbol{x}) K_h\left(\frac{z_i - z}{h}\right)} \tag{4}$$

where $\mathbb{I}(\bullet)$ is the indicator function. Therefore, this approach relies on the selection of a product kernel function $K_h(\bullet)$ and an optimal bandwidth parameter $h > 0$ selected using any choice method. In the empirical application presented in Section 3, we adopt the data-driven selection approach developed by Bădin et al. (2010) that is based on the Least Squares Cross-Validation (LSCV) method (Hall et al., 2004; Li & Racine, 2004). The recent work of Bădin et al. (2019) presents practical aspects of bandwidth selection for conditional efficiency measures.

2.3 | Spatialized frontier models and the new SpDEA approach

The importance of considering spatial aspects of data when estimating efficiency scores is widely recognized, as failure to do so may result in biased or inefficient estimates. However, it is only in recent years that studies have begun to explore methods for incorporating information about the structure of spatial heterogeneity and/or spatial dependence of geographically distributed productive units when estimating technical efficiency. In response to this need, a new class of spatial frontier models has emerged in the productivity analysis literature. Recent surveys by Orea and Álvarez (2022) and Ayouba (2023) highlight the importance of this approach, which emphasizes the significance of uneven performance in space (spatial heterogeneity), geographical interconnections (spatial dependence), and spillover effects. By considering spatial heterogeneity and dependence, these models can produce more accurate and reliable estimates of efficiency, providing a valuable means of understanding the complexities of regional productivity and identifying opportunities for improvement.

In the nonparametric literature, there are two approaches that try to account for the presence of autocorrelated data in the field of productive efficiency: the *peers of location* method proposed by Vidoli and Canello (2016), which can be considered a "pure" spatial production frontier method, and the *conditional frontier* approach of Daraio and Simar (2005, 2007), which uses a more indirect approach to control for spatial effects.

In Vidoli and Canello (2016), unlike DEA where each evaluated productive unit is compared to a convex linear combination of the DMUs in the sample, each DMU is evaluated only against a subset of productive units - the potential competitors - that are located in close proximity of the evaluated DMU. The crucial point of this methodology is the identification of the set of local peers of each DMU, the authors proposing to use an empirical semi-variogram to estimate the optimal distance in terms of spatial autocorrelation which is used to select the set of spatial peers. Auteri et al. (2019) proposed to use a less rigid geostatistical tool, the Skater algorithm, to form clusters and then find the peers of a given DMU. Complementarily, Fusco et al. (2018) propose to use a k-nearest neighbors approach to locate the peers of each DMU.

On the other hand, in trying to disentangle the intrinsic performance of productive units from the influence of conditioning factors that may affect efficiency, Daraio and Simar (2005, 2007) propose the inclusion of external contextual variables. These factors should explain the heterogeneity between productive units through administrative, political, socio-economic, environmental, and other variables that control for differences among DMUs. Particularly, the spatial dependence could be controlled introducing variables directly related to the specific geographical location of the units such as geographical coordinates of DMUs or spatial lag variables. In this paper, we choose the second option, using a conditional efficiency framework where the conditioning variables for each location are the weighted averages (convex combinations) of the output per worker and the capital per worker of neighboring regions. The details are provided below.

In our empirical application, we aim to construct a spatial frontier model for the EU28 regions. To achieve this, we start by formulating a regional production function $Y_{it} = F(K_{it}, L_{it})$, where Y_{it} represents the aggregate output in region i at year t, K_{it} denotes the respective stock of physical capital, and L_{it} is the employed labor force. As is usual in the productivity growth literature at the macro level, the constant returns-to-scale (CRS) assumption will be used (at the national level see, for example, Färe et al., 1994, Kumar & Russell, 2002, Henderson & Russell, 2005, or Los & Timmer, 2005; at regional level see, among others, by Enflo & Hjertstrand, 2009, Filippetti & Peyrache, 2015, or Beugelsdijk et al., 2018). Hence, in the production frontier, the output per worker (labor productivity) ratio ($q_{it} = Y_{it}/L_{it}$) is a function only of the capital per worker ('capital deepening') input ($x_{it} = K_{it}/L_{it}$), $q_{it} = f(x_{it})$. In essence, this production model implies that, for any given per worker input of physical capital, there is an efficient level of output per worker. In other words, under this simple unconditional frontier approach, the difference in the level of labor productivity between an efficient region and another inefficient one is due only to two components, the distance of the inefficient region to the frontier, and the different capital intensity (level of capital stock per unit of labor) between the two regions. However, this difference does not depend on the location of the compared regions.

To introduce the spatial dimension in the DEA model, accounting for the spatial dependence observed in the regional data, a spatial lag of the output per worker ratio (a weighted average of the level of productivity in neighboring regions), $z_{1i} = \sum_{j=1}^{n} w_{ij} q_j = \boldsymbol{w}_i' \boldsymbol{q}$, and a spatial lag of the capital-labor ratio (a weighted average of capital intensity in nearby regions), $z_{2i} = \sum_{j=1}^{n} w_{ij} x_j = \boldsymbol{w}_i' \boldsymbol{x}$, are used as components of external conditioning factors of the vector $\boldsymbol{z}$ in the DEA efficiency approach.[2] The

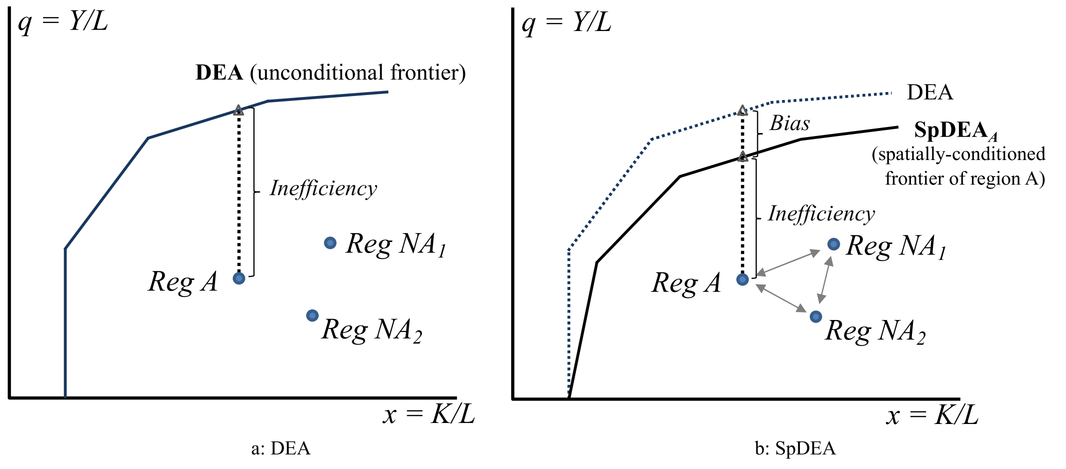

FIGURE 1 Unconditional and spatially conditioned DEA regional frontiers.

spatial weights, w_{ij}, capture the spatial interaction between regions i and j (these elements are known a priori and satisfy the conditions $w_{ij} \geq 0$, $w_{ii} = 0$, and $\boldsymbol{w}_i'\mathbf{1} = \sum_{j=1}^{n} w_{ij} = 1$), and the resulting n by n matrix $\mathbf{W}$ with elements w_{ij} describes the connectivity of the n regions.[3]

Figure 1 provides a visual summary of the construction of the regional frontiers for the unconditional (Figure 1a) and spatially conditioned DEA (Figure 1b) approaches, representing labor productivity against capital intensity for the case of different regional economies.

We will focus on region A, denoted as *Reg A*. Assume the existence of geographical externalities among region A and their neighboring regions, denoted as *Reg* NA_1 and *Reg* NA_2. The presence of these geographical externalities would imply changes in the shape of the production frontier, changing regional technical efficiency in our target region, *Reg A*. Effectively, we would expect spatial dependence among the efficiencies of *Reg A*, and its neighbors *Reg* NA_1 and *Reg* NA_2, and different behaviors of the spatial factors could produce distinct shapes of the efficient frontier for *Reg A*.

In the standard approach (unconditional DEA, Figure 1a), the measure of inefficiency for region A, represented by the vertical distance from *Reg A* to the production frontier, does not consider the presence of spatial externalities derived from the existence of geographical spillovers among region A and their neighboring regions A_1 and A_2. However, in the SpDEA approach (spatially conditioned case, Figure 1b), apart from the unconditional frontier, there exists a specific frontier for the case of region A that depends on the level of a set of spatial factors related to the location of this region (contextual variables that influence the frontier level of region A). In Figure 1b, the measure of inefficiency for region A considers the presence of geographical externalities produced from its neighboring regions. Now, the vertical distance from *Reg A* to the spatially conditioned frontier is showing a different estimation of the inefficiency than that obtained for the unconditional case.

What Figure 1 reveals is that, under the presence of geographical externalities, the estimation of the regional inefficiencies with the standard DEA for region A are downward-biased. However, while a standard DEA productivity analysis might yield results that would not reflect the impact of geographical externalities, our SpDEA approach would try to reduce the potential for such bias.

Thus, we may postulate a consensus on the thesis that from a hypothetical regional empirical application using the DEA and the SpDEA models, the expectation would be an increase in the scores. Less agreement can be claimed for the role played by the existing spatial dependence on the shape of the distribution of the estimated technical efficiencies, and this is where one of the contributions of our method lies.

3 | EMPIRICAL APPLICATION

In this section, region-level technical efficiency in the EU-28 over the 2000–2014 period will be analyzed applying the standard DEA and the new SpDEA model proposed in the previous section in order to account for the presence of heterogeneous conditions in the form of geographical externalities.[4] Our hypothesis is that the proposed frontier model, in trying to capture the role of contextual variables in the form of geographical externalities that are beyond the control of regional or even national authorities, can help to explain efficiency differentials between regions that contribute to the generation of different regional productivity growth trajectories.

3.1 | Data

In our application, we concentrate on analyzing the performance of 263 NUTS-2 regions across the 28 European Union countries (before 2020), excluding the overseas territories of Finland, France, Portugal and Spain. The data used in the empirical application were taken from the Cambridge Econometrics' European Regional Database (ERD) 2016 release that contains complete yearly information for the period 1990–2014 at the regional NUTS-2 classification of the European Union.[5]

From the ERD, the following variables were calculated or estimated.

- Regional output (Y), measured as gross value added -GVA-in each region in constant 2005 purchasing power standards -PPS- terms. The original GVA at constant prices time series (measured in constant 2005 million euros) were adjusted for price differences across countries and over the time with country-specific PPS's.
- Regional labor (L), measured as total employment in each region in thousands of persons.
- Regional PPS levels of GVA were divided by the total number of workers, and thus a series of real GVA per worker (q) was calculated for each region.
- Gross fixed capital formation (I), measured in million euros (2005 prices).
- To obtain estimates of regional real net capital stocks (K), the Perpetual Inventory Method (PIM) was employed using yearly real gross fixed capital formation (I) series based on the equation $K_{it} = I_{it} + (1-\delta)K_{i,t-1}$. The investment series for the period 1990–2000 were accumulated to obtain an initial estimate of regional level of real net capital stock in the year 2000, $K_{i0} = K_{i,2000}$. Thereafter, all the stocks in this year were benchmarked to the aggregate estimates of the national capital stocks in the year 2000 (measured in constant 2005 million euros) reported in Penn World Tables 8.1 (Feenstra et al., 2015) using the share of each region in the national capital stock in the year 2000 obtained with the 1990–2000 investment-accumulated regional capital stock estimations. The PIM formula was then applied to calculate the real net capital stock estimates for the period 2001–2014, using an annual depreciation rate of 5%.
- Regional series of per worker capital (x) were obtained dividing regional levels of physical capital stocks by the total number of workers of each region.

3.2 | Results

In this subsection the empirical results of the analysis for the case of the 263 EU-28 NUTS-2 regions will be presented and discussed.

TABLE 1 Descriptive statistics on DEA and SpDEA efficiency scores.

	2000		2007		2014	
	DEA	SpDEA	DEA	SpDEA	DEA	SpDEA
Mean	0.424	0.586	0.464	0.584	0.454	0.576
Std. Dev.	0.150	0.223	0.130	0.144	0.117	0.130
Min.	0.123	0.123	0.219	0.219	0.236	0.245
Median	0.431	0.627	0.438	0.607	0.438	0.602
Max.	1	1	1	1	1	1
Pearson's correlation (p-value)	0.823 (0.000)		0.653 (0.000)		0.537 (0.000)	
Spearman's rank correlation (p-value)	0.835 (0.000)		0.684 (0.000)		0.493 (0.000)	

3.2.1 | Comparing DEA and SpDEA results

First, we will compare the regional technical efficiency scores estimated with both the baseline DEA model and the SpDEA model proposed in this paper.[6]

Aspatial DEA scores are estimated using (2) and, drawing on (6), the SpDEA technical efficiency of an observed region ($x_{it} = K_{it}/L_{it}$, $q_{it} = Y_{it}/L_{it}$) facing external conditions $\left(z_{1i} = \boldsymbol{w}_i'\boldsymbol{q}, z_{2i} = \boldsymbol{w}_i'\boldsymbol{x}\right)$ is given by:

$$\hat{\lambda}_{SpDEA}\left(x_i, q_i | \boldsymbol{w}_i'\boldsymbol{q}, \boldsymbol{w}_i'\boldsymbol{x}\right) = sup\left\{\lambda > 0 \mid \hat{S}_{Q|X,Z_1Z_2}\left(\lambda q_i | X \leq x_i, Z_1 = \boldsymbol{w}_i'\boldsymbol{q}, Z_2 = \boldsymbol{w}_i'\boldsymbol{x}\right) > 0\right\} \quad (9)$$

Table 1 provides summary statistics of the efficiency estimates of both models in the years 2000, 2007 and 2014. Starting from the descriptive statistics, we observe that the mean of both unconditional and conditional efficiencies generally increased or remained stable from 2000 to 2007.[7] Furthermore, comparing the average regional efficiencies for each year, we find that the SpDEA estimates are consistently higher than the DEA estimates, with increases of 38.2%, 25.9%, and 26.9% in 2000, 2007, and 2014, respectively. These findings suggest that a significant portion of regional technical inefficiencies in the EU can be attributed to interregional spillovers in the production process.

On the other hand, the standard deviation of both DEA and SpDEA efficiency scores provide straightforward indicators of sigma-convergence for European regions over the period 2000–2014. The results shows that the standard deviation of DEA (SpDEA) efficiency scores declined from 0.150 (0.223) in the year 2000 to 0.117 (0.130) in the year 2014. The ridgelines plots of Figure 2 corroborate graphically this result of narrowing of the distributions of efficiency scores over time.

Complementarily, if the correlations of the efficiency scores for the DEA and SpDEA models are compared, it can be appreciated that Pearson's and Spearman's rank coefficients vary from about 0.823 to 0.835 in 2000 to 0.537 and 0.493, respectively, in 2014. Hence, although there has been a reduction in the dispersion of the efficiency scores (sigma-convergence), the relationship between the two scores variables has decreased during this period of time.

In summary, the descriptive results of Table 1 show that the impact of neighbors on a region's level of efficiency can be considerable, indicating the importance of incorporating spatial effects into the DEA methodology so widely used in efficiency analyses. From a graphical point of view, Figure 2 complements this result. Therefore, the density traces provide new information, indicating the shape of the distributions for the technical efficiency estimates: the scores show bimodal distributions—twin peaks - in the case of the SpDEA model, especially in the years 2000 and 2014, pointing to the presence of two clusters ('efficiency clubs') of regional scores.[8]

FIGURE 2 Ridgeline plots of DEA and SpDEA efficiency scores.

The statistical property of multimodality in the distribution of the SpDEA scores is corroborated in Table 2, where we provide the results of the Excess-Mass (EM) tests for multimodality (Ameijeiras-Alonso et al., 2019) of DEA and SpDEA efficiency scores. From the results reported in this table, we have determined the number of modes in the different distributions of DEA and SpDEA efficiency scores. Note that the number of modes in the distributions of DEA efficiency scores is one for all the years, while the distributions of the SpDEA estimates present two modes in the years 2000 and 2014, showing the polarization of the European regional technical efficiency. This result highlights, again, the bias of the baseline DEA in the construction of the production-frontier.

3.2.2 | Marginal local impact of z's on the regions' efficiency scores

To assess the direction of the influence of the spatial variables $z_{1i} = \boldsymbol{w}_i'\boldsymbol{q}$ and $z_{2i} = \boldsymbol{w}_i'\boldsymbol{x}$ on the production process, we investigate the partial regression plots of the spatial conditioned to unconditional

TABLE 2 Tests for multimodality of DEA and SpDEA efficiency scores.

Year	DEA	SpDEA
2000	EM = 0.027	EM = 0.073
	p-value = 0.838	*p*-value <0.000
	Est. loc. = 0.427	Est. loc. = 0.369, 0.667
2007	EM = 0.031	EM = 0.033
	p-value = 0.546	*p*-value = 0.558
	Est. loc. = 0.423	Est. loc. = 0.653
2014	EM = 0.031	EM = 0.069
	p-value = 0.600	*p*-value <0.000
	Est. loc. = 0.423	Est. loc. = 0.410,0.613

Note: Excess-Mass -EM- test [H_0: One Mode; H_1: More Than One Mode] and estimated location of modes -Est. loc-. Bootstrapped *p*-values are used, with 5000 bootstrap replications.

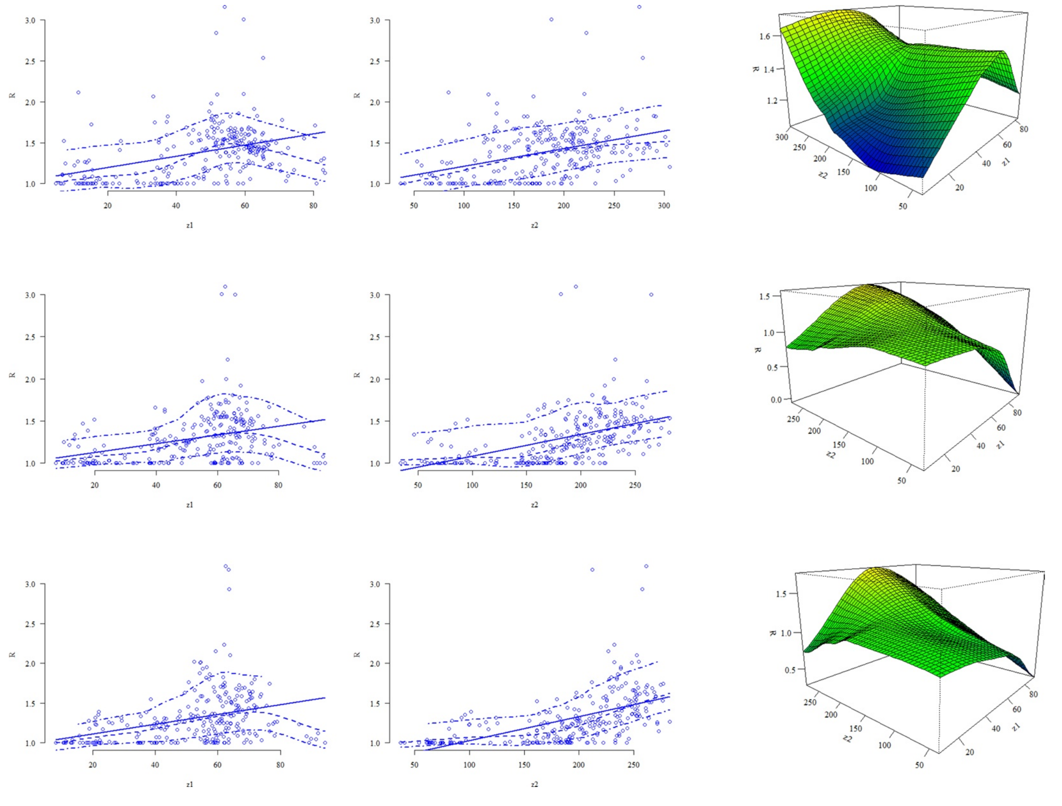

FIGURE 3 Local impact of spatial variables $z_{1i} = \boldsymbol{w}_i'\boldsymbol{q}$ and $z_{2i} = \boldsymbol{w}_i'\boldsymbol{x}$ on the production process (years 2000, 2007 and 2014, by rows).

efficiency score ratio, $\hat{R}_i$, with respect to the variables z_{1_i} and z_{2i}. Figure 3 displays the results of these graphs for the years 2000, 2007, and 2014.

Focusing on the effect of z_1 (spatial lag of the output per worker ratio) on the shift of the frontier, we can see an approximately inverted 'U' shape form in the local regression lines. Furthermore, we might have some shift of the efficient boundary when the real GVA per worker ratio of the neighbors

TABLE 3 Effects of spatial external variables in the SpDEA model.

	2000	2007	2014
	p-value	*p*-value	*p*-value
$z_{1i} = \boldsymbol{w}_i'\boldsymbol{q}$	0.09*	<0.001***	<0.001***
$z_{2i} = \boldsymbol{w}_i'\boldsymbol{x}$	0.41	<0.001***	<0.001***

Note: Kernel regression significance tests (type I test with IID bootstrap: 1000 replications).

increases, but after a certain threshold in the level of the output per labor variable, a decreasing effect is revealed. Hence, regional labor productivity appears to play a significant role in accelerating regional technological change (shifts in the production frontier). The reported threshold effect seems to confirm the findings of Ramajo et al. (2017) for 120 regions in EU-9 over the 1995–2007 period, where the GVA spatial lag variable also has an inverted U-shaped non-linear impact on the performance of regions, interpreting this finding as a differential effect of interregional spillovers depending on the size of the neighboring economies.

On the other hand, there is a more linear shape for the effect of z_2 (spatial lag of the per worker capital input) on the regions' efficient boundary in the 3 years considered. When the capital-labor ratio grows, shifts of the EU regional frontier occur, the *K*/*L* variable acting as an accelerator of the technical change. This evidence points to the favorable effect of the level of capital intensity on the production process via technology change.

Additionally, in Table 3, we test the statistical significance of each conditioning variable using the kernel-weighted local linear least squares non-parametric method and the significance test discussed in Section 2. As can be seen through the corresponding *p*-values, the two spatial external variables, z_1 and z_2, are statistically significant in years 2007 and 2014, but only z_1 is weakly significant in the year 2000. Therefore, it seems that the contextual variables used in this work had a smaller impact on the input-output space at the beginning of the period under analysis.

3.2.3 | Mapping and exploratory spatial analysis of the SpDEA efficiency scores

Considering the above results with respect to the statistical significance of the spatial variables, the technical efficiency estimates derived from the SpDEA model are mapped and described. In the first place, the first row of Figure 4 shows the percentile maps of the regional SpDEA efficiency scores in 2000, 2007 and 2014, and in the second row the Moran's scatterplots and *I* statistics are shown; finally, the local Moran's I (LISA cluster maps) are shown in the third row.

Looking at the percentile maps, there is clear positive spatial dependence between the yearly regional efficiency scores. This spatial autocorrelation is verified by the scatterplots and the Moran's *I* statistics (0.763 in 2000, 0.616 in 2007, and 0.541 in 2014, with pseudo *p*-values less than 0.001 in all cases). It is important to highlight that the temporal evolution of the three Moran's scatterplots shows a final pattern (year 2014) where almost all the regions are located in the first (High-High efficiency scores) and third (Low-Low efficiency scores) quadrants. This pattern was corroborated by the corresponding LISA maps, where the existence of two spatial clusters for the EU-28 regions can be identified: the first cluster (the High-High zone) is characterized by highly efficient regions surrounded by regions with similar scores, while the second cluster (the Low-Low zone) includes more inefficient regions. Geographically, the first cluster is located in the United Kingdom and in central and north-western Europe (‘*core*’), while the second group includes the regions in eastern Europe and the most southern continental regions of Spain and Portugal (‘*periphery*’). Thus, the twin-peaks (bimodality)

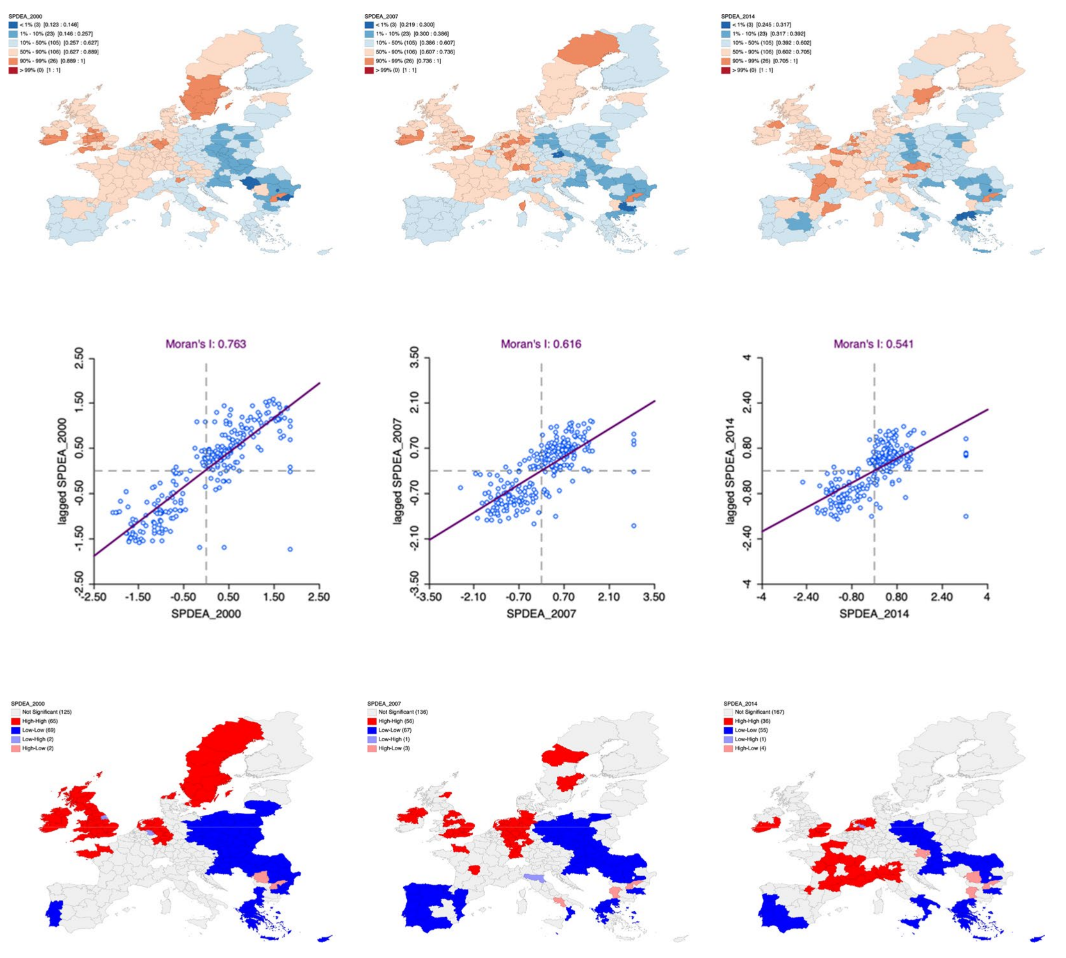

FIGURE 4 Percentile maps, Moran's scatterplots and LISA cluster maps of SpDEA efficiency scores (years 2000, 2007 and 2014, by columns).

property noted earlier for the density distributions of SpDEA efficiency scores is associated with spatial heterogeneity in the form of two different core-periphery spatial efficiency clubs.

3.2.4 | Discussion of the results

Our results identify geographical externalities as partially responsible for changes in technical efficiencies in the European regions. First, regional technological change can be induced by the output per worker ratio of the neighboring regions (through the different microfoundations of agglomeration economies that could generate production cost savings[9]). Secondly, the proximity or density of high capital per worker inducing technology change is another mechanism through which the regional technical efficiencies change.

Also, our estimates of regional technical efficiencies reveal that, over the analyzed period, the efficiency scores have bimodal distributions. This is a new finding for the European regional case that could be related to the detected multimodal distribution in European regional productivity. For example, the kernel distributions of the regional efficiencies estimated by Rogge (2019) across NUTS-2 regions in the EU for the years 2000 and 2011 demonstrate that the distribution of regional productivity was multimodal but did not detect the existence of multimodality in the distribution of efficiency scores. The failure to detect this characteristic could be the result of the aspatial nonparametric production frontier technique used by the author to estimate the scores.

Additionally, in general, the regional distribution of labor productivity in Europe has shown the existence of two convergence productivity clubs. These clubs maintain internal relationships characterized by a core-periphery geographical pattern (Benito & Ezcurra, 2005; De Siano & D'Uva, 2006; Villaverde & Maza, 2008; Basile, 2009; Escribá-Pérez & Murgui-García, 2018). From our results, it could be hypothesized that this polarization in regional productivity could be driven primarily by a bimodal technical efficiency distribution; in this case, geographical externalities would be an important determinant of the technical efficiency. This would imply that the two clubs of European regions in terms of regional convergence have their counterfactual basis in two categories of European regions that are showing different technical efficiencies.

3.2.5 | Technical efficiency and regional competitiveness

In this subsection, we perform an exploratory analysis to uncover the underlying mechanisms driving the observed spatial heterogeneity. Through this exploration, we establish potential contributions and implications of these mechanisms in relation to prior empirical studies.

Efficiency and competitiveness are two key concepts that are closely related in the context of regional development. Regional technical efficiency is a key determinant of regional competitiveness. To try to explain the polarization found in our estimated SpDEA scores, we will carry out a descriptive analysis that attempts to correlate the estimated regional technical efficiency for the year 2014 with the European Regional Competitiveness Index (RCI) - and their components - proposed by the European Commission to measure the major factors of competitiveness of the regions within the EU.[10] These correlations can provide important insights into the factors that are driving regional productivity and competitiveness, and can help them to design policies that promote economic growth and development at the regional level.

The RCI is a measure of the performance of the NUTS-2 regions within the EU; the 11 dimensions of competitiveness, pillars, each one capturing concepts related to regional productivity and long-term economic growth. As argued by Dijkstra et al. (2011, p. 4), *"This Index (…) assumes that the focus in less developed economies should be different than in intermediate or highly developed economies. Whereas less developed economies need to ensure that their basic transport infrastructure, basic education and health care services are of a good quality, highly developed regions should be more concerned about their business sophistication and their use of technology and innovation"*. As result, this logic can be expected to compare the correlations between every one of the two identified spatial clusters for the EU-28 regions (the High-High zone characterized by highly efficient regions surrounded by regions with similar scores, and the second Low-Low cluster of regions including more inefficient regions and surrounded by regions with similar scores) and the different pillars.

The correlation that links our 2014 SpDEA efficiency scores for the case of the 263 regions with the RCI variable yields a statistically significant value of 0.557 (p-value <0.001). This is a sign that our measure of technical efficiency is related in part to the general competitiveness of the European regional economies. The correlations were also examined by subgrouping the sample into the two clusters of regions. Therefore, Table 4 shows the correlation results relating our SpDEA efficiency scores for both High-High regions and Low-Low regions in the year 2014 with the different components of the RCI.[11]

TABLE 4 Correlations between SpDEA efficiency scores of 2014 and the different components of the Regional Competitiveness Index (RCI) by the two regional subgroups.

	High-high regions (147 regions)	Low-low regions (89 regions)
RCI	0.088	0.171
(1) Institutions	−0.031	**0,309*** **
(2) Macroeconomic stability	−0.111	**0,301*** **
(3) Infrastructure	**0,148***	0.110
(4) Health	0.098	**−0,280*** **
(5) Basic education	**−0,165** **	**0,376*** **
(6) Higher education and training	0.052	0.050
(7) Labor market efficiency	−0.023	**0,274*** **
(8) Market size	**0,190** **	0.043
(9) Technological readiness	−0.049	0.141
(10) Business sophistication	**0,180** **	**−0,205** **
(11) Innovation	0.028	0.062

Note: Correlations in bold type are statistically significant

*** *p*-value <0.01, ** *p*-value <0.05, * *p*-value <0.1.

There were significant differences in the correlations for the two subgroups. For the case of the High-High regional cluster, the components "Market Size," "Business Sophistication" and "Infrastructure" are positively related to SpDEA scores, while "Basic Education" is negatively related with SpDEA scores.[12] On the other hand, for the Low-Low cluster, the components of "Basic Education," "Institutions," "Macroeconomic Stability" and "Labor Market Efficiency" are positively related to SpDEA scores, whereas "Health" and "Business Sophistication" are negatively related with SpDEA scores.[13]

To further explore the polarization that the European regional economies are evidencing, and according to the detected correlations, some general comments are provided. First, the empirical evidence obtained for the two subgroups show that the detected bimodal distribution in the European regional efficiency seems to be driven by different factors. As a general conclusion, the results concerning the correlations substantially confirm the predicted expectations (Dijkstra et al., 2011). The correlations indicate the existence of a more complex pattern, with different dimensions of competitiveness operating within every regional cluster. The findings from the previous exploratory analysis have been confirmed by several recent studies that have examined the determinants of regional productivity through confirmatory analysis. Some of this research suggests that spatial heterogeneities lead regional productivity to react distinctively to alterations in the factors influencing it. Thus, in regions marked by low-quality institutions, the relevance of physical capital seems restricted (Rodríguez-Pose & Ganau, 2022).[14] Another example illustrating the different effects generated by the drivers of the regional productivity for the case of the two clusters detected by our study is related to Cainelli et al. (2022), who show that the growth advantage linked with membership in international business groups is observed exclusively in regional environments with low-quality institutions.[15] Finally, the presence of different factors influencing regional productivity within each regional cluster is the observation that labor productivity growth is more pronounced in regions with a superior endowment of social capital (Ganau & Rodríguez-Pose, 2023), anticipating a varied response among regions categorized into the two regional subgroups. [16] These studies confirm the existence of a more complex pattern regarding the determinants of regional productivity. Their results are linked to various dimensions of factors influencing regional productivity operating within each regional cluster identified in the distribution of efficiency scores estimated through SpDEA. Therefore, the novelty of our paper lies in both the methodological integration of space within the standard DEA (as a conventional DEA analysis would conceal the uncovered results) and the identification of bimodal distributions in the efficiency scores of European regions using SpDEA (these distributions allow us to detect the existence of determinants of regional productivity related to the presence of interregional externalities).

Secondly, these substantial regional disparities also highlight the need for different and more tailored European regional policies when addressing measures to improve regional efficiency as a way to increase regional productivity, an important channel to make regions more competitive.

Therefore, these partial correlations could be indicating, in part, the underlying mechanisms responsible for the existence of low or high levels of efficiency within every one of these two groups of European regions. Technical efficiency is the key way that regions can raise productivity and so, to increase their per-capita income in route to a path toward achieving regional economic convergence.

Additionally, although it has not been noted previously in the literature, the determinants of the existence of twin peaks in the efficiency of the European regions would be important indicators of the causes of the differences in productivity between the US and Europe. Indeed, the European productivity slowdown relative to the United States of America has been an important focus for applied researchers. Among the diverse studies, a great deal of attention has been paid to the slower emergence in Europe of the so called "knowledge-based economy" (Griffith et al., 2006; Van Ark et al., 2008). The efficiencies derived from the Information and Communication Technologies within a regional economy (and their associated geographical externalities) could be a key element to stimulate the European productivity (Miller & Atkinson, 2014). Nevertheless, this could be a simplistic expression of a more complex problem: while the 'knowledge economy' is a fundamental explanation for the European productivity gap, the effects from other causes could be also operating (see, for example, a review by Ortega-Argilés, 2012). Hence, at the regional level, other factors, such as institutional qualities, market size or the regional innovative capacity may matter for regional productivity (see Pike et al., 2017).

3.2.6 | Policy implications

From our empirical results, future policy challenges related to European productivity should attempt to foster efficiency through economic policies that focus on the different requirements of the two types of regions. The conclusions derived from the empirical results obtained in the case of the European regions can provide the analytical foundations when it comes to explaining the failures of European regional policies to achieve more competitive regions: policies suitable for High-High regions may not work for Low-Low regions. In other words, our research shows that, depending on the type of region, different policy recommendations could emerge about how to improve regional technical efficiency as a way to increase regional competitiveness. This would imply addressing the geographical externalities of regional economic policies beyond both a general policy recommendation for the total group of regions under analysis and the single region objectives they pursue. Comprehensive policies that cover the operational field of interregional spillovers could be in the design of measures ready to break the existence of dichotomies within European regional economies through the existence of twin peaks in regional productivity, regional convergence and, as shown in this paper, regional efficiency. To address the issue of interregional spillovers and the existence of dichotomies within European regional economies, it is essential to develop comprehensive policies that cover all aspects of this phenomenon. Such policies should aim to break down the twin peaks in regional productivity, regional convergence, and, as demonstrated in this paper, enhance regional efficiency. By implementing measures designed to address these issues, we can work toward creating a more competitive and cohesive European regional economy.

4 | SUMMARY AND CONCLUDING REMARKS

In this paper, it is argued that the standard DEA method used to estimate efficiencies with spatial data fails to acknowledge that the efficiencies of many DMUs are correlated with the performances of the neighboring units. Consequently, in order to avoid the corresponding bias, a more appropriate model should capture the role of geographical externalities when focusing on spatial data. Since the external factors introduced in our proposed conditional DEA model are the spatial lags of the input and the output of the underlying aggregate production function, the proposed frontier specification can be called a Spatial DEA model. In order to illustrate the relevance of the proposal, the results from the traditional DEA and the SpDEA models have been considered for the case of the analysis of 263 NUTS-2 European regions.

The empirical application for the 2000 (economic boom), 2007 (pre-financial crisis) and 2014 (post-financial crisis) years have revealed that, after controlling for geographical externalities, the average EU-28 regional efficiency has increased by 16.2 p. p., 12.0 p. p. and 12.2 p. p., respectively. This result shows that in standard DEA applications, without the appropriate spatial corrections (as we have argued in the second section), inefficiency is overestimated, providing evidence that a non-negligible part of regional technical efficiency in the European Union is related to interregional spillovers in the production process. Ignoring spatial dependence underestimates efficiency gains in regions with highly significant geographical spillovers. This finding could challenge the appropriateness of policies that have, for example, under-allocated public resources to such regions, or could raise doubts about the estimation of the level of 'budgetary waste' in EU Member States' spending applying standard DEA to countries or regions' production levels of public services (Saulnier, 2020).

On the other hand, the dispersion of the spatially conditioned scores has declined over time (the standard deviation declined from 0.223 in the year 2000 to 0.130 in the year 2014), showing that EU-28 regional technical efficiency has sigma-converged during the period under analysis, 2000–2014. Furthermore, the SpDEA scores show bimodal distributions - twin peaks - especially in the years 2007 and 2014, pointing to the presence of two spatial clusters (or efficiency clubs) of regional scores. Although this distributional pattern has been identified previously in the literature in terms of both European regional productivity and convergence of per capita incomes, the detection of twin peaks by means the DEA technique in the regional efficiency within EU is a new finding. As a general conclusion, the SpDEA model can help to explain efficiency differentials between European regions that contribute to the generation of different regional productivity growth trajectories.

An implication of our results is that the technical efficiency in the European regions may be underestimated in studies that use the standard DEA model in order to try to explain the low levels of regional productivity, hiding the peaks and the troughs in the spatial distribution of regional technical efficiencies. The exploratory analysis of our research suggests that, depending on the allocation to the two types of regions, different policy recommendations emerge about how to improve regional technical efficiency as a way to increase regional competitiveness. Consequently, the decision reached in terms of policy recommendations (European financial support for investments and reforms) in order to encourage technical efficiency should contemplate the different requirements of the regions according to their belonging to the two types of spatial efficiency clubs.

A future research priority should be the investigation of the underlying mechanisms that are maintaining these two spatial clusters of efficiency in the European regions. Another future development would be to extend our analysis to the sectoral level, studying the efficiencies of different industries. A more high-level consequence of the spatial heterogeneity analysis carried out in this paper is that contemporaneous inter-country or multi-regional macroeconomic models need to consider the existence of spatial spillovers between cross-units, and therefore explicitly include spatial aspects in its formulation. These spatially based theoretical models, and their empirical counterparts, will be more useful in order to better understand issues such as global income inequality, productivity convergence between countries, or differential regional impacts of any innovation within the EU or any other geographic area, to cite just a few examples.

ACKNOWLEDGEMENTS

Julián Ramajo and Miguel Á. Márquez acknowledge the financial support by Spanish Ministerio de Ciencia e Innovación, Grant No. PID2019-109687GB-I00, funded by MCIN/AEI/10.13039/501100011033. This work was supported by the Ministry of Science and Innovation of Spain under-Grant PID2019-109687GB-I00.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

STATEMENTS AND DECLARATIONS

Authors disclose financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

ORCID

Julián Ramajo https://orcid.org/0000-0002-3156-8315
Miguel A. Márquez https://orcid.org/0000-0003-0784-5025
Geoffrey J. D. Hewings https://orcid.org/0000-0003-2560-3273

ENDNOTES

[1] Further as Proost and Thisse (2019) have noted, there is an important link between microeconomic spatial forces/foundations and their contributions to generating and sustaining spatial inequalities. The role and importance of spatial spillovers associated with flows of goods and services, people, and innovations create dynamics that can have profound implications for the performance of the macroeconomy.

[2] The inclusion of the $\boldsymbol{w'q}$ and $\boldsymbol{w'x}$ variables can be rationalized from a theoretical point of view if we think in a production function such as $Y = AF(K,L)$ where A (which represents technological progress, incorporating aspects of knowledge generation, innovation, commuting, sectoral linkages or level of human capital at the NUTS-2 regional level) depends on the spatial lags of both endogenous and exogeneous variables. See, for example, Kock, 2010, where a spatially augmented Solow parametric model with physical capital externalities and spatial externalities is developed and estimated using 204 European NUTS-2 regions over the 1977–2000 period.

[3] In our empirical application, we used a weight matrix defined as $\mathbf{W} = \{w_{ij} = 1 \; if \; i \neq j \text{ and } j \in N_5(i); w_{ij} = 0 \text{ otherwise}\}$, $N_5(i)$ being the set of 5-nearest neighbors to i. We chose five neighbors as cut-off point to the connections of i because 5 was the median of neighbors between the spatial units of our sample data. In addition to the 5-nearest neighbor matrix, we also used the traditional (first order) contiguity matrix and the inverse distance matrix as alternative weighting matrices, and the results were almost identical to the presented in Section 3, These complementary results are available upon request to the corresponding author.

[4] The *R* code used in the calculations is available from the corresponding author.

[5] The primary source of the ERD is the Eurostat's REGIO database, supplemented with the European Commission's AMECO database. The 2016 release of ERD uses the NUTS 2010 regional classification.

[6] We are conscious that there are two issues, noisy data and endogeneity problems, not addressed in the estimation of regional efficiencies. A robust order-*m* approach could be used to obtain estimates more resistant to outliers and extreme values (Cazals et al., 2002), and the endogeneity issue could be addressed by using the procedure proposed in Simar et al. (2016). These extensions will be the objective of future consideration.

[7] Specifically, the DEA efficiency score increased from 0.424 in 2000 to 0.464 in 2007, while the SpDEA efficiency score increased from 0.586 in 2000 to 0.584 in 2007. However, both scores decreased from 2007 to 2014, although the changes were relatively small, with the DEA score dropping from 0.464 in 2007 to 0.454 in 2014, and the SpDEA score dropping from 0.584 in 2007 to 0.576 in 2014.

[8] See the mapping and exploratory spatial analysis paragraph below to find out the membership of the regions to each of the efficiency clubs.

[9] These micro-foundations were described by Marshall (1920) and Rosenthal and Strange (2004); for a review, see Cohen and Morrison (2009) and Puga (2010).

[10] The most recent version of the RCI ("RCI 2.0–2022 edition") has been published in March 2023 (Dijkstra et al., 2023). In our calculations we have used the "RCI—2019 edition" data for the year 2016, which is based on more than 70 comparable indicators which are mostly referred to the period 2013–2014, our year of reference (2014) for the ongoing descriptive correlation analysis. The RCI indices were launched in 2010 and published every three years; therefore, there is no information for the RCIs that permits comparison with the SpDEA scores of the years 2000 and 2007.

[11] Annoni et al. (2019) uses a similar approach to examine regional resilience in the European Union. Therefore, first spatial clusters within the European Union are identified by employing exploratory spatial data analysis, and then a spatial econometric model is employed to estimate the determinants of regional growth in both regimes, using the RCI dimensions as explanatory variables as well as other factors of growth.

[12] With respect to the "Market Size" pillar, this indicator embraces characteristics such as regional GDP and population. The relation between market size and productivity has been recognized by different authors (for example, Bernhardt, 1981; Epifani & Gancia, 2006). Antonelli et al. (2011) and Siller et al. (2021) have focused on the relevance of knowledge and technology-intensive sectors (business sophistication). The contribution of "Infrastructure" to the productivity across European regions has been studied by numerous studies (see, for example, Basile, 2009; Del Bo & Florio, 2012).

[13] "Basic Education" seems to be an important pillar when studying regional technical efficiency within the low group of regions. "Institutions" is an indicator of the quality of government that reflects the perceptions of the citizens with respect to corruption, quality and impartial allocation in their public-sector services. The important role of institutions for economic performance (Rodríguez-Pose, 2010; Rodríguez-Pose & Storper, 2006) has been highlighted recently by Rodríguez-Pose and Ganau (2019) and Agostino et al. (2020) when analyzing the determinants of regional labor productivity trajectories in Europe. The general economic climate ("Macroeconomic stability") exerts a positive influence in the technical efficiency of these regions. Other authors have examined the influence of "Labor Market Efficiency" on European productivity (Cyrek & Fura, 2019; Escribá-Pérez & Murgui-García, 2018).

[14] This also seems to be indicated from an exploratory perspective in Table 4 for the Low-Low regions, where the lack of significance of the correlation between the SpDEA efficiency scores and the "Infrastructure" component is demonstrated. At the same time, for the case of the High-High regions, the correlation is significant and positive.

[15] As membership in a business group could influence the efficiency of the labor market, affecting factors such as labor mobility and adaptability to changes, the findings from Cainelli et al. (2022) support the significant and positive correlation estimated in the Low-Low regions subgroup in Table 4 for the 'Labor Market Efficiency' component. This correlation was found to be non-significant for the other subgroup of High-High regions.

[16] Once again, in Table 4, the responses differ for the two subgroups: the correlation between the SpDEA efficiency scores and the components of the RCI related to social capital ("Institutions" and "Macroeconomic Stability") is only significant in the case of the Low-Low regions.

REFERENCES

Agostino, M., Di Tommaso, M. R., Nifo, A., Rubini, L., & Trivieri, F. (2020). Institutional quality and firms' productivity in European regions. *Regional Studies*, *54*(9), 1275–1288. https://doi.org/10.1080/00343404.2020.1712689

Alcidi, C. (2019). Economic integration and income convergence in the EU. *Intereconomics*, *54*(1), 5–11. https://doi.org/10.1007/s10272-019-0783-6

Alvaredo, F., Chancel, L., Piketty, T. S., & Zucman, G. (Eds.) (2018)., *World inequality report 2018*. Belknap Press.

Ameijeiras-Alonso, J., Crujeiras, R. M., & Rodríguez-Casal, A. (2019). Mode testing, critical bandwidth and excess mass. *Test*, *28*(3), 900–919. https://doi.org/10.1007/s11749-018-0611-5

Annoni, P., de Dominicis, L., & Khabirpour, N. (2019). Location matters: A spatial econometric analysis of regional resilience in the European union. *Growth and Change*, *50*(3), 834–855. https://doi.org/10.1111/grow.12311

Antonelli, C., Patrucco, P. P., & Quatraro, F. (2011). Productivity growth and pecuniary knowledge externalities: An empirical analysis of agglomeration economies in European regions. *Economic Geography*, *87*(1), 23–50. https://doi.org/10.1111/j.1944-8287.2010.01104.x

Auteri, M., Guccio, C., Pammolli, F., Pignataro, G., & Vidoli, F. (2019). Spatial heterogeneity in non-parametric efficiency: An application to Italian hospitals. *Social Science and Medicine*, *239*, 112544. https://doi.org/10.1016/j.socscimed.2019.112544

Ayouba, K. (2023). Spatial dependence in production Frontier models. *Journal of Productivity Analysis*, *60*(1), 21–36. https://doi.org/10.1007/s11123-023-00670-7

Bădin, L., Daraio, C., & Simar, L. (2010). Optimal bandwidth selection for conditional efficiency measures: A data-driven approach. *European Journal of Operational Research*, *201*(2), 633–640. https://doi.org/10.1016/j.ejor.2009.03.038

Bădin, L., Daraio, C., & Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model? *European Journal of Operational Research*, *223*(3), 818–833. https://doi.org/10.1016/j.ejor.2012.06.028

Bădin, L., Daraio, C., & Simar, L. (2019). A bootstrap approach for bandwidth selection in estimating conditional efficiency measures. *European Journal of Operational Research*, *277*(2), 784–797. https://doi.org/10.1016/j.ejor.2019.02.054

Badinger, H., Müller, W., & Tondl, G. (2004). Regional convergence in the European union, 1985-1999: A spatial dynamic panel analysis. *Regional Studies*, *38*(3), 241–253. https://doi.org/10.1080/003434042000211105

Badunenko, O., Henderson, D. J., & Russell, R. R. (2013). Polarization of the worldwide distribution of productivity. *Journal of Productivity Analysis*, *40*(2), 153–171. https://doi.org/10.1007/s11123-012-0328-5

Badunenko, O., Henderson, D. J., & Zelenyuk, V. (2018). The productivity of nations. In E. Grifell-Tatjé, C. A. Knox Lovell, & R. C. Sickles (Eds.), *The oxford handbook of productivity analysis*. Oxford University Press.

Basile, R. (2009). Productivity polarization across regions in Europe: The role of nonlinearities and spatial dependence. *International Regional Science Review*, *32*(1), 92–115. https://doi.org/10.1177/0160017608326944

Benito, J. M., & Ezcurra, R. (2005). Spatial disparities in productivity and industry mix: The case of the European regions. *European Urban and Regional Studies*, *12*(2), 177–194. https://doi.org/10.1177/0969776405053743

Bernhardt, I. (1981). Sources of productivity differences among Canadian manufacturing industries. *The Review of Economics and Statistics*, *63*(4), 503–512. https://doi.org/10.2307/1935845

Beugelsdijk, S., Klasing, M. J., & Milionis, P. (2018). Regional economic development in Europe: The role of total factor productivity. *Regional Studies*, *52*(4), 461–476. https://doi.org/10.1080/00343404.2017.1334118

Blanchet, T., Chancel, L., & Gethin, A. (2019). How unequal is Europe? Evidence from distributional national accounts, 1980-2017. WID.world Working Paper, No. 2019/06.

Bruno, R. L., Douarin, E., Korosteleva, J., & Radosevic, S. (2019). *Determinants of productivity gap in the European union: A multilevel perspective*. Institute of Labor Economics (IZA) Discussion Papers. No. 12542.

Cainelli, G., Ganau, R., & Giunta, A. (2022). Business groups, institutions, and firm performance. *Industrial and Corporate Change*, *31*(1), 215–233. https://doi.org/10.1093/icc/dtab046

Cazals, C., Florens, J. P., & Simar, L. (2002). Nonparametric Frontier estimation: A robust approach. *Journal of Econometrics*, *106*(1), 1–25. https://doi.org/10.1016/s0304-4076(01)00080-x

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Charnes, A., Cooper, W. W., & Rhodes, E. (1981). Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. *Management Science*, *27*(6), 668–697. https://doi.org/10.1287/mnsc.27.6.668

Cohen, J. P., Coughlin, C. C., & Paul, C. J. M. (2019). Agglomeration, productivity and regional growth: Production theory approaches. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories: Revised and extended* (2nd ed., pp. 104–123). Edward Elgar Publishing.

Cohen, J. P., & Morrison, P. C. J. (2009). Agglomeration, productivity and regional growth: Production theory approaches. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories*. Edward Elgar. Cheltenham 101–117.

Cosci, S., & Mirra, L. (2018). A spatial analysis of growth and convergence in Italian provinces: The role of road infrastructure. *Regional Studies*, *52*(4), 516–527. https://doi.org/10.1080/00343404.2017.1334117

Cyrek, M., & Fura, B. (2019). Employment for sustainable development: Sectoral efficiencies in EU countries. *Social Indicators Research*, *143*(1), 277–318. https://doi.org/10.1007/s11205-018-1970-8

Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric Frontier models: A probabilistic approach. *Journal of Productivity Analysis*, *24*(1), 93–121. https://doi.org/10.1007/s11123-005-3042-8

Daraio, C., & Simar, L. (2007). Conditional nonparametric Frontier models for convex and non-convex technologies: A unifying approach. *Journal of Productivity Analysis*, *28*(1), 13–32. https://doi.org/10.1007/s11123-007-0049-3

Davis, M. A., Fisher, J. D. M., & Whited, T. M. (2014). Macroeconomic implications of agglomeration. *Econometrica*, *82*(2), 731–764.

De Groot, H., Poot, J., & Smit, M. J. (2016). Which agglomeration externalities matter most and why. *Journal of Economic Surveys*, *30*(4), 756–782. https://doi.org/10.1111/joes.12112

Del Bo, C. F., & Florio, M. (2012). Infrastructure and growth in a spatial framework: Evidence from the EU regions. *European Planning Studies*, *20*(8), 1393–1414. https://doi.org/10.1080/09654313.2012.680587

De Siano, R., & D'Uva, M. (2006). Club convergence in European regions. *Applied Economics Letters*, *13*(9), 569–574. https://doi.org/10.1080/13504850600733473

De Witte, K., & Kortelainen, M. (2013). What explains performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics*, *45*(17), 2401–2412. https://doi.org/10.1080/00036846.2012.665602

Dijkstra, L., Annoni, P., & Kozovska, K. (2011). *A new regional competitiveness index: Theory, methods and findings*. European Commission, Directorate-General for Regional Policy. WP 02/2011.

Dijkstra, L., Papadimitriou, E., Cabeza-Martínez, B., de Dominicis, L., & Kovacic, M. (2023). *EU regional competitiveness index 2.0 - 2022 edition*. European Commission, Directorate-General for Regional and Urban Policy. WP 01/2023.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2063–2117). Elsevier B. V.

Emrouznejad, A., & Yang, G. L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, *61*(1), 4–8. https://doi.org/10.1016/j.seps.2017.01.008

Enflo, K., & Hjertstrand, P. (2009). Relative sources of European regional productivity convergence: A bootstrap frontier approach. *Regional Studies*, *43*(5), 643–659. https://doi.org/10.1080/00343400701874198

Epifani, P., & Gancia, G. (2006). Increasing returns, imperfect competition, and factor prices. *The Review of Economics and Statistics*, *88*(4), 583–598. https://doi.org/10.1162/rest.88.4.583

Escribá-Pérez, F. J., & Murgui-García, M. J. (2018). Technology catching-up and regulation in European regions. *Journal of Productivity Analysis*, *49*(1), 95–109. https://doi.org/10.1007/s11123-017-0524-4

Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, *84*(1), 66–83.

Färe, R., & Lovell, C. K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, *19*(1), 150–162. https://doi.org/10.1016/0022-0531(78)90060-1

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn world table. *The American Economic Review*, *105*(10), 3150–3182. https://doi.org/10.1257/aer.20130954

Filippetti, A., & Peyrache, A. (2015). Labor productivity and technology gap in European regions: A conditional Frontier approach. *Regional Studies*, *49*(4), 532–554. https://doi.org/10.1080/00343404.2013.799768

Fracasso, A., & Vittucci-Marzetti, G. (2018). Estimating dynamic localization economies: The inadvertent success of the specialization index and the location quotient. *Regional Studies*, *52*(1), 119–132. https://doi.org/10.1080/00343404.2017.1281388

Fritsch, M., & Kublina, S. (2018). Related variety, unrelated variety and regional growth: The role of absorptive capacity and entrepreneurship. *Regional Studies*, *52*(10), 1360–1371. https://doi.org/10.1080/00343404.2017.1388914

Fritsch, M., & Slavtchev, V. (2011). Determinants of the efficiency of regional innovation systems. *Regional Studies*, *45*(7), 905–918. https://doi.org/10.1080/00343400802251494

Fusco, E., Vidoli, F., & Sahoo, B. K. (2018). Spatial heterogeneity in composite indicator: A methodological proposal. *Omega*, *77*, 1–14. https://doi.org/10.1016/j.omega.2017.04.007

Ganau, R., & Rodríguez-Pose, A. (2023). Firm-level productivity growth returns of social capital: Evidence from Western Europe. *Journal of Regional Science*, *63*(3), 529–551. https://doi.org/10.1111/jors.12636

Gobillon, L., & Magnac, T. (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. *The Review of Economics and Statistics*, *98*(3), 535–551. https://doi.org/10.1162/rest_a_00537

Griffith, R., Huergo, E., Mairesse, J., & Peters, B. (2006). Innovation and productivity across four European countries. *Oxford Review of Economic Policy*, *22*(4), 483–498. https://doi.org/10.1093/oxrep/grj028

Haelermans, C., & De Witte, K. (2012). The role of innovations in secondary school performance: Evidence from a conditional efficiency model. *European Journal of Operational Research*, *223*(2), 541–549. https://doi.org/10.1016/j.ejor.2012.06.030

Hall, P., Racine, J., & Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association*, *99*(468), 1015–1026. https://doi.org/10.1198/016214504000000548

Henderson, D. J., & Russell, R. R. (2005). Human capital and convergence: A production-frontier approach. *International Economic Review*, *46*(4), 1167–1205. https://doi.org/10.1111/j.1468-2354.2005.00364.x

Jeong, S. O., Park, B. U., & Simar, L. (2010). Nonparametric conditional efficiency measures: Asymptotic properties. *Annals of Operations Research*, *173*(1), 105–122. https://doi.org/10.1007/s10479-008-0359-5

Johnson, P., & Papageorgiou, C. (2020). What remains of cross-country convergence? *Journal of Economic Literature*, *58*(1), 129–175. https://doi.org/10.1257/jel.20181207

Kerr, W. R., & Kominers, S. D. (2015). Agglomerative forces and cluster shapes. *The Review of Economics and Statistics*, *97*(4), 877–899. https://doi.org/10.1162/rest_a_00471

Kock, W. (2010). Growth and spatial dependence in Europe. In A. Páez, J. Gallo, R. N. Buliung, & S. Dall'erba (Eds.), *Progress in spatial analysis: Methods and applications*. Springer Science and Business Media.

Kubis, A., & Schneider, L. (2016). Regional migration, growth and convergence: A spatial dynamic panel model of Germany. *Regional Studies*, *50*(11), 1789–1803. https://doi.org/10.1080/00343404.2015.1059932

Kumar, S., & Russell, R. (2002). Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *The American Economic Review*, *92*(3), 527–548. https://doi.org/10.1257/00028280260136381

Li, Q., & Racine, J. (2004). Cross-validated local linear nonparametric regression. *Statistica Sinica*, *14*(2), 485–512.

Los, B., & Timmer, M. P. (2005). The 'appropriate technology' explanation of productivity growth differentials: An empirical approach. *Journal of Development Economics*, *77*(2), 517–531. https://doi.org/10.1016/j.jdeveco.2004.04.001

Marshall, A. (1920). *Principles of economics*. Macmillan.

Maudos, J., Pastor, J. M., & Serrano, L. (2000). Efficiency and productive specialization: An application to the Spanish regions. *Regional Studies*, *34*(9), 829–842. https://doi.org/10.1080/00343400020002967

Miller, B., and Atkinson, R. D. (2014) *Raising European productivity growth through ICT*. ITIF, June 2014. https://doi.org/10.2139/ssrn.3079844

Moradi-Motlag, A., & Emrouznejad, A. (2022). The origins and development of statistical approaches in non-parametric Frontier models: A survey of the first two decades of scholarly literature (1998–2020). *Annals of Operations Research*, *318*(1), 713–741. https://doi.org/10.1007/s10479-022-04659-7

Orea, L., & Álvarez, I. C. (2022). Production economics in spatial analysis. In S. C. Ray, R. Chambers, & S. Kumbhakar (Eds.), *Handbook of production economics* (pp. 1379–1409). Springer. https://doi.org/10.1007/978-981-10-3450-3_35-1

Ortega-Argilés, R. (2012). The transatlantic productivity gap: A survey of the main causes. *Journal of Economic Surveys*, *26*(3), 395–419. https://doi.org/10.1111/j.1467-6419.2012.00725.x

Parr, J. B. (2002). Agglomeration economies: Ambiguities and confusions. *Environment & Planning A*, *34*(4), 717–731. https://doi.org/10.1068/a34106

Phelps, N. A., Fallon, R. J., & Williams, C. L. (2001). Small firms, borrowed size and the urban-rural shift. *Regional Studies*, *35*(7), 613–624. https://doi.org/10.1080/00343400120075885

Pike, A., Rodríguez-Pose, A., & Tomaney, J. (2017). Shifting horizons in local and regional development. *Regional Studies*, *51*(1), 46–57. https://doi.org/10.1080/00343404.2016.1158802

Proost, S., & Thisse, J. F. (2019). What can be learned from spatial economics? *Journal of Economic Literature*, *57*(3), 575–643. https://doi.org/10.1257/jel.20181414

Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, *50*(1), 203–219. https://doi.org/10.1111/j.1467-9787.2009.00657.x

Racine, J. (1997). Consistent significance testing for nonparametric regression. *Journal of Business and Economic Statistics*, *15*(3), 369–378. https://doi.org/10.1080/07350015.1997.10524714

Racine, J., Hart, J. D., & Li, Q. (2006). Testing the significance of categorical predictor variables in nonparametric regression models. *Econometric Reviews*, *25*(4), 523–544. https://doi.org/10.1080/07474930600972590

Ramajo, J., Cordero, J. M., & Márquez, M. A. (2017). European regional efficiency and geographical externalities: A spatial nonparametric Frontier analysis. *Journal of Geographical Systems*, *19*(4), 319–348. https://doi.org/10.1007/s10109-017-0249-y

Ramajo, J., & Hewings, G. J. D. (2018). Modelling regional productivity performance across Western Europe. *Regional Studies*, *52*(10), 1372–1387. https://doi.org/10.1080/00343404.2017.1390219

Ramajo, J., Márquez, M. A., Hewings, G. J. D., & Salinas, M. M. (2008). Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions? *European Economic Review*, *52*(3), 551–567. https://doi.org/10.1016/j.euroecorev.2007.05.006

Ray, S. C. (2020). Data envelopment analysis: A nonparametric method of production analysis. In S. Ray, R. Chambers, & S. Kumbhakar (Eds.), *Handbook of production economics*. Springer. https://doi.org/10.1007/978-981-10-3450-3_24-1

Rodríguez-Pose, A. (2010). Economic geographers and the limelight: Institutions and policy in the world development report 2009. *Economic Geography*, *86*(4), 361–370. https://doi.org/10.1111/j.1944-8287.2010.01094.x

Rodríguez-Pose, A., & Ganau, R. (2019). *Institutions and the productivity challenge for European regions (No. 116)*. Directorate General for Economic and Financial Affairs (ECFIN).

Rodríguez-Pose, A., & Ganau, R. (2022). Institutions and the productivity challenge for European regions. *Journal of Economic Geography*, *22*(1), 1–25. https://doi.org/10.1093/jeg/lbab003

Rodríguez-Pose, A., & Storper, M. (2006). Better rules or stronger communities? On the social foundations of institutional change and its economic effects. *Economic Geography*, *82*(1), 1–25. https://doi.org/10.1111/j.1944-8287.2006.tb00286.x

Rogge, N. (2019). Regional productivity growth in the EU since 2000: Something is better than nothing. *Empirical Economics*, *56*(2), 423–444. https://doi.org/10.1007/s00181-017-1366-7

Rosenthal, S. S., & Strange, W. C. (2004). *Evidence on the nature and sources of agglomeration economies*. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2119–2171). Elsevier.

Saulnier, J. (2020). *Improving the quality of public spending in Europe: Budgetary 'waste rates' in EU Member States*. Directorate General for Parliamentary Research Service (EPRS).

Siller, M., Schatzer, T., Walde, J., & Tappeiner, G. (2021). What drives total factor productivity growth? An examination of spillover effects. *Regional Studies*, *55*(6), 1129–1139. https://doi.org/10.1080/00343404.2020.1869199

Simar, L., Vanhems, A., & Van Keilegom, I. (2016). Unobserved heterogeneity and endogeneity in nonparametric Frontier estimation. *Journal of Econometrics*, *190*(2), 360–373. https://doi.org/10.1016/j.jeconom.2015.06.015

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, *39*(3), 312–320. https://doi.org/10.2307/1926047

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, *49*(2), 326–365. https://doi.org/10.1257/jel.49.2.326

Tsekeris, T., & Papaioannou, S. (2018). Regional determinants of technical efficiency: Evidence from the Greek economy. *Regional Studies*, *52*(10), 1398–1409. https://doi.org/10.1080/00343404.2017.1390312

Tsvetkova, A., Ahrend, R., Martins, J. O., Lembcke, A. C., Knutsson, P., Jong, D., & Terzidis, N. (2020). The spatial dimension of productivity: Connecting the dots across industries, firms and places. OECD Regional Development Working Papers, No. 2020/01. 1–65. https://doi.org/10.1787/ba5edb47-en

Van Ark, B., O'Mahoney, M., & Timmer, M. P. (2008). The productivity gap between Europe and the United States: Trends and causes. *The Journal of Economic Perspectives*, *22*(1), 25–44. https://doi.org/10.1257/jep.22.1.25

Vidoli, F., & Canello, J. (2016). Controlling for spatial heterogeneity in nonparametric efficiency models: An empirical proposal. *European Journal of Operational Research*, *249*(2), 771–783. https://doi.org/10.1016/j.ejor.2015.10.050

Villaverde, J., & Maza, A. (2008). Productivity convergence in the European regions, 1980–2003: A sectoral and spatial approach. *Applied Economics*, *40*(10), 1299–1313. https://doi.org/10.1080/00036840600771361

Vines, D., & Wills, S. (2018). The rebuilding macroeconomic theory project: An analytical assessment. *Oxford Review of Economic Policy*, *34*(1–2), 1–42. https://doi.org/10.1093/oxrep/grx062

Widuto, A. (2019) *Regional inequalities in the EU, European parliamentary research service*, , May 2019, European Parliament.

How to cite this article: Ramajo, J., Márquez, M. A., & Hewings, G. J. D. (2024). Addressing spatial dependence when estimating technical efficiency: A spatialized data envelopment analysis of regional productive performance in the European Union. *Growth and Change*, e12711. https://doi.org/10.1111/grow.12711

APPENDIX 1 THE CONDITIONAL FRONTIER NONPARAMETRIC METHODOLOGY

To analyze the influence of contextual factors on the production function or the inefficiency distribution, it is necessary to reformulate the optimization problem associated with the baseline DEA model. This can be achieved by defining the production process using the probabilistic formulation developed by Cazals et al. (2002) and Daraio and Simar (2005, 2007).

The production process can be described by the probability measure of the joint random vector ($\boldsymbol{X}$,$\boldsymbol{Q}$) denoted by $H_{X,Q}$ ($\boldsymbol{x}$,$\boldsymbol{q}$):

$$H_{X,Q}(\boldsymbol{x},\boldsymbol{q}) = \mathrm{Prob}(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Q} \geq \boldsymbol{q}) \tag{A1}$$

This probability function can be further decomposed as follows:

$$H_{X,Q}(\boldsymbol{x},\boldsymbol{q}) = \mathrm{Prob}(\boldsymbol{Q} \geq \boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x})\mathrm{Prob}(\boldsymbol{X} \leq \boldsymbol{x}) = S_{Q|X}(\boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x})F_X(\boldsymbol{x}) \tag{A2}$$

where $S_{Q|X}$ ($\boldsymbol{q}$|$\boldsymbol{X} \leq \boldsymbol{x}$) represents the conditional survival function of $\boldsymbol{Q}$ and $F_X(\boldsymbol{x})$ is the cumulative distribution function of $\boldsymbol{X}$.

The probabilistic formulation of the production process is able to handle the presence of observed heterogeneity in the form of k factors $\boldsymbol{z} = (z_1, z_2, \ldots, z_k) \in \mathbb{R}_+^k$ that might have an influence on the production process. For this purpose, a conditional efficiency approach is developed to account for such variables in the frontier estimation by conditioning the production process to a given value of the environmental random vector $\boldsymbol{Z}$, $\boldsymbol{Z} = z$. Consequently, efficiency estimates are determined by inputs, outputs and external variables. The joint probability function can be extended and decomposed as follows:

$$\begin{aligned} H_{X,Q|Z}(\boldsymbol{x},\boldsymbol{q}|\boldsymbol{z}) &= \mathrm{Pr}(\boldsymbol{Q} \geq \boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z})\mathrm{Pr}(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}) \\ &= S_{Q|X,Z}(\boldsymbol{q}|\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z})F_{X|Z}(\boldsymbol{x}|\boldsymbol{Z} = \boldsymbol{z}) \end{aligned} \tag{A3}$$

The function $H_{X,Q|Z}$ ($\boldsymbol{x}$,$\boldsymbol{q}$|z) represents the probability of a DMU operating at level ($\boldsymbol{x}$,$\boldsymbol{q}$) being dominated by other DMUs facing the same conditions z that might influence the production.

By means of the conditional frontier approach, the direction of the effect of the external factors on the production process can be evaluated by comparing conditional with unconditional efficiency scores (Bădin et al., 2012). Using the ratio of efficiency estimates, it will be possible to analyze the impact that $\boldsymbol{Z}$ may have on the shape of the efficient boundary (shifts in the frontier), in such a way that the conditioning variables can affect the range of attainable values of the input-output space.

More specifically, if the z's are continuous variables, then the influence of the environmental variables on the distribution of the efficiency scores can be quantified by computing the graph of the ratios of conditional to unconditional scores against the components of $\boldsymbol{z_i}$ and its smoothed nonparametric regression function over the sample of n observations $\{(\boldsymbol{x}_i, \boldsymbol{q}_i, z_i) | i = 1, 2, \ldots, n\}$:

$$\hat{R}\left(\boldsymbol{x_i}, \boldsymbol{q_i} | z_i\right) = \frac{\hat{\lambda}_{cDEA}\left(\boldsymbol{x_i}, \boldsymbol{q_i} | z_i\right)}{\hat{\lambda}_{DEA}\left(\boldsymbol{x_i}, \boldsymbol{q_i}\right)} \tag{A4}$$

By generating these partial regression plots, we can easily inspect the direction of the effect of the environmental variables upon the efficiency ratios: in our output-oriented conditional model, an increasing regression curve indicates that the corresponding variable in $\boldsymbol{z}$ is favorable to efficiency (the conditional frontier moves to the unconditional one when the specific variable z increases), whereas a decreasing line will denote an negative effect on the production process (the conditional efficient boundary moves away from the unconditional frontier).

Like-wise, it is also possible to investigate the statistical significance of the explanatory variables $\boldsymbol{z}$ explaining the variations of $\hat{R}\left(\boldsymbol{x_i}, \boldsymbol{q}_i | \boldsymbol{z}_i\right)$. Following Badin et al. (2010), and Jeong et al. (2010), local linear least squares can be used for the regression estimation, followed by the nonparametric regression significance test proposed by Li and Racine (2004). Specifically, the significance of each variable is tested using bootstrap resampling tests as proposed by Racine (1997) and Racine et al. (2006); this approach can be interpreted as the nonparametric equivalent of standard t-ratio tests in ordinary least squares regression (De Witte & Kortelainen, 2013; Haelermans & De Witte, 2012).