Journal of Regional Research
Investigaciones Regionales

Nº 64 - 2026/1

Articles

ISSN: 1695-7253

Our Staff

Investigaciones Regionales
Journal of Regional Research

ISSN: 1695-7253 **E-ISSN: 2340-2717**

Facultad de Ciencias Económicas, Empresariales y Turismo
Universidad de Alcalá.
Plaza de la Victoria, 2, 28802, Alcalá de Henares, Madrod
Teléfono: +34 91 885 42 09
E-mail: investig.regionales@aecr.org
www.investigacionesregionales.org

Journal of Regional Research
Investigaciones Regionales

Volume 2026/1 - Issue 64

Articles

ISSN: 1695-7253 E-ISSN: 2340-2717
investig.regionales@aecr.org
AECR - Asociación Española de Ciencia Regional
www.aecr.org
España - Spain

Investigaciones Regionales – Journal of Regional Research is included in the following databases:

- ESCI – Emerging Sources Citation Index (Web of Science, Clarivate Analytics)
- SCOPUS
- RePEc (Research Papers in Economics)
- Recyt (Repositorio Español de Ciencia y Tecnología de la FECTYT – Fundación Española para la Ciencia y la Tecnología)
- DOAJ (Directory of Open Access Journals)
- Redalyc (Red de Revistas Científicas de América Latina y el Caribe, España y Portugal)
- Latindex (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, Caribe, España y Portugal)
- EconLit (American Economic Association (AEA), Estados Unidos) – Econlit with Full Text (EBSCO Publishing)
- Dialnet (Universidad de La Rioja, España)
- CARHUS Plus+ sistema de evaluación de revistas científicas de los ámbitos de las Ciencias Sociales y las Humanidades que se publican a nivel local, nacional e internacional (AGAUR)
- Cabell's Directory (Cabell Publishing, Inc.)
- Fuente Académica Plus
- ProQuest (ABI/INFORM Complete; ABI/INFORM Global; Professional ABI/INFORM Complete; ProQuest Central; ProQuest 5000 International; ProQuest 5000)
- e_Bu@h – Biblioteca Digital Universidad de Alcalá
- Road – Directory of Oper Access Scholary Resources

Articles

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 5-16
https://doi.org/10.38191/iirr-jorr.24.064

Articles

A test of the relationship between the Pareto exponent and sample size

Rafael González-Val, Fernando Sanz-Gracia***

Received: 01 April 2024
Accepted: 01 October 2024

Abstract:

This paper uses un-truncated city population data from three countries—the United States, Spain and Italy—to empirically test Proposition 1 put forth by Eeckhout (2004 *American Economic Review*, 94: 1429–1451). Eeckhout's hypothesis was that the estimate of the Pareto exponent in a standard Zipf regression decreases with sample size, if the underlying city size distribution is lognormal. Using rolling sample regressions, we find that this proposition is only valid once we enter the lognormal body of the distribution; for the Pareto-distributed upper-tail, the estimated exponent does not vary with sample size.

Keywords: City size distribution; Zipf's law; Pareto exponent; Pareto distribution; lognormal distribution; rolling sample regressions.
JEL Classification: C12; R11; R12.

Una prueba de la relación entre el exponente de Pareto y el tamaño muestral

Resumen:

Este documento utiliza datos de población de ciudades sin restricciones de tamaño de tres países—Estados Unidos, España e Italia—para poner a prueba empíricamente la Proposición 1 presentada por Eeckhout (2004 *American Economic Review*, 94: 1429–1451). La hipótesis de Eeckhout era que la estimación del exponente de Pareto en una regresión Zipf estándar disminuye con el tamaño de la muestra, si la distribución del tamaño de las ciudades subyacente es lognormal. Utilizando regresiones de muestra móvil, encontramos que esta proposición solo es válida una vez que entramos en la parte central lognormal de la distribución; para la cola superior distribuida siguiendo una función de Pareto, el exponente estimado no varía con el tamaño muestral.

Palabras clave: Distribución del tamaño de la ciudad; ley de Zipf; exponente de Pareto; distribución de Pareto; distribución lognormal; regresiones de muestra móvil.
Clasificación JEL: C12; R11; R12.

1. Introduction

Zipf's law is an empirical regularity that has received significant attention in the urban economics and geographic literature. It establishes a linear and stable relationship between the rank and size (population) of cities and is considered to be a reflection of a steady-state situation.

* Universidad de Zaragoza, Facultad de Economía y Empresa, Departamento de Análisis Económico, Zaragoza. España. / Institut d'Economia de Barcelona (IEB), Facultat d'Economia i Empresa, Barcelona. España. rafaelg@unizar.es
** Universidad de Zaragoza, Facultad de Economía y Empresa, Departamento de Análisis Económico, Zaragoza. España. fsanz@unizar.es
Corresponding author: rafaelg@unizar.es

The significance of Zipf's law is difficult to overstate. First, due to its broad applicability across numerous fields, it can be applied to virtually any quantitative phenomenon. For instance, it has been used to study the size distribution of the number of victims in armed conflicts (González-Val, 2016), the frequency of musical notes in famous compositions (Zanette, 2006), the magnitude of migratory movements (Clemente et al., 2011), and, most famously, the frequency of different words in Joyce's Ulysses (Zipf, 1949). Second, it has clear theoretical ties to urban growth theory, as shown by Gabaix (1999). Third, there is an almost esoteric quality to the empirical regularity derived from Zipf's law, which is known as the rank-size rule: the k-th largest city is exactly one k-th the size of the largest city. This empirical regularity has fascinated urban geographers since Auerbach's (1913) seminal work. Finally, in the context of urban planning and demography, Cristelli et al. (2012, p. 7) argued that for Zipf's law to hold, the urban system must be integrated and the sample must be coherent in the sense of being the "result of some kind of optimization in growth processes or of an optimal self-organization mechanism." That situation implies that the urban system shares a common language, culture, history, and set of rules.

In an influential article, Eeckhout (2004) stimulated an academic debate about what distribution better fits city size distributions. He justified the use of un-truncated city size data, showing that the underlying statistical distribution has strong implications for the fulfilment of Zipf's law. Traditionally, due to data limitations, most studies have considered only the largest cities. However, Eeckhout (2004) demonstrated the statistical importance of considering both large and small cities because truncated samples lead to biased results. In particular, he found that the Zipf exponent declines systematically as the sample size increases if the underlying distribution is lognormal rather than Pareto (i.e., Proposition 1 in Eeckhout (2004)).

However, that proposition has remained largely untested (a few exceptions include Luckstead and Devadoss (2014) and Peña and Sanz-Gracia (2021)), probably because the reaction of the literature was a search for the best distribution to fit un-truncated data and the lognormal distribution was soon replaced by other more convoluted ones. The current consensus is that the best city size distribution may be a mixed distribution, separating the lognormal body of the distribution from the upper Pareto tail (Giesen et al., 2010; Ioannides and Skouras, 2013; Puente-Ajovín et al., 2020). However, the method for defining the Pareto upper-tail is still subject to debate (Fazio and Modica, 2015; Schluter, 2021). Therefore, the possible relationship between Zipf's law and the sample size for both the upper-tail and the whole distribution is still a relevant issue that we aim to clarify in this study, which constitutes our primary innovative contribution to the existing literature.

The remainder of this paper is organised as follows. Section 2 briefly reviews the literature. Section 3 presents the data and the methodology that we used. Section 4 describes our main results, and Section 5 concludes our work.

2. Literature review

There are excellent surveys on this topic that comprehensively review the literature on city size distribution and Zipf's law up to the time of their publication. In this regard, we can chronologically cite Nitsch (2005), Cottineau (2017), and Arshad et al. (2018). Accordingly, this paragraph lists a selection of recent or relatively recent papers (from 2018 onwards) that address the topic of Zipf's law for cities: Arshad et al. (2019), Hackmann and Klarl (2020), Düben and Krause (2021), Fernholz and Kramer (2024), and González-Val et al. (2024).

Our contribution to the existing literature lies in explicitly testing Proposition 1 of Eeckhout (2004), which establishes a direct relationship between the behavior of the Pareto exponent and sample size. While many studies have traditionally estimated the Pareto exponent for different sample sizes (e.g., Eeckhout (2004), González-Val (2010)), analyses that examine variations in the parametric estimate of the Pareto exponent across all possible sample sizes are less common. To the best of our knowledge, only Peng (2010), Fazio and Modica (2015), and Peña and Sanz-Gracia (2021) have used a recursive procedure (i.e., rolling sample regressions) to test the relationship between the Pareto exponent and sample size. An alternative branch of the literature estimates the local Pareto exponent by sample size using nonparametric methods (Ioannides and Overman, 2003; Luckstead and Devadoss, 2014).

3. Data and methods

We used settlement size data from the decennial censuses of three countries: the United States (US; 2000 and 2010 censuses), Spain (2001 and 2011 censuses) and Italy (2001 and 2011 censuses). The data were obtained from the national official statistical services and included un-truncated city population data without any size restrictions.

For the US, our sample for the year 2000 is the same as that used by Eeckhout (2004). The spatial units are what the US Census Bureau calls 'places,' which include both incorporated places (i.e., administrative cities legally incorporated under the laws of their respective states) and Census Designated Places (i.e., a concentration of population, housing and commercial structures that is identifiable by name, but is not within an incorporated place). We considered 25,358 places in the US in the 2000 dataset and 29,461 places in the 2010 dataset.

The geographical unit of reference in Spain and Italy is the municipality. Municipalities are the smallest spatial units (local governments); they are the administratively defined "legal" cities, comprising the whole territory and population of both countries. For Italy, the number of cities by period is 8,100 municipalities in 2001 and 8,081 in 2011. For Spain, our samples include 8,108 municipalities in 2001 and 8,074 in 2011.

The standard Zipf regression equation, including the correction introduced by Gabaix and Ibragimov (2011), is as follows:

$$\ln\left(R-\frac{1}{2}\right)=b-a\ln S+\varepsilon, \tag{1}$$

where R is the empirically observed rank (1 for the largest city, 2 for the second largest, and so on), S is city size (population) and a is the Pareto exponent. If $\hat{a}=1$, Zipf's law holds, meaning that, ordered from largest to smallest, the size of the second city is half that of the first, the size of the third is a third of the first, and so on. In any case, a is interpreted as a measure of the degree of inequality in the city size distribution: large (small) values correspond to more equal (unequal) city sizes. Standard errors are calculated by applying Gabaix and Ioannides's (2004) correction: $\mathrm{GI\,s.e.}=\hat{a}\cdot(2/N)^{1/2}$, where N is the sample size. We use these corrected standard errors to calculate the confidence bands of $\hat{a}$ at the 95% confidence level.

Proposition 1 of Eeckhout (2004, p. 1442) reads as follows: "If the underlying distribution is the lognormal distribution, then the estimate of the parameter $\hat{a}$ of the Pareto distribution is increasing in the truncation city size and decreasing in the truncated sample population." That is, adding increasingly smaller cities to a sample should result in growing inequality in the distribution. This statement is what we sought to test empirically. To do so, we employed rolling sample regressions (Peng, 2010; Fazio and Modica, 2015; Peña and Sanz-Gracia, 2021): we incorporated cities one by one into the sample until we attained the smallest urban unit. We began with the largest two cities, and each time we calculated the estimated Pareto exponent.

4. Results

Figures 1, 2 and 3 show our results for the three countries we studied[1]. In each figure, a graph is shown for the whole city size distribution in any of the two years considered, as well as an additional graph zooming in the upper-tail of the distribution. To define the upper-tail, we followed the procedure set forth by Clauset et al. (2009), which is specifically designed to select an optimal truncation point[2]. Note that this cut-off is only used to delimit the number of estimates shown in the upper-tail graphs. As a robustness check, we estimated alternative thresholds using the procedure of Beirlant et al. (1996) recommended by

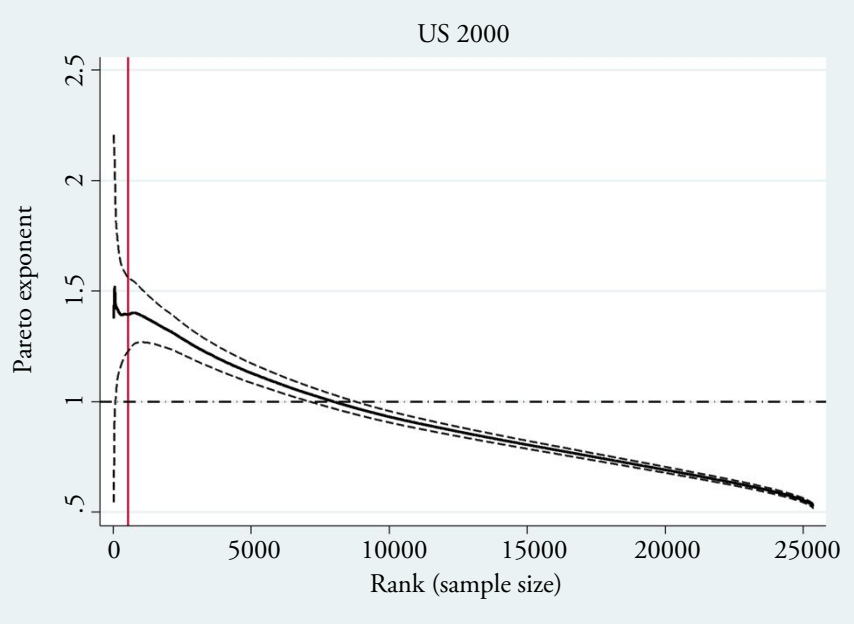
[1] Figure 1 is similar to Figure 2 of Fazio and Modica (2015).

[2] For a review of the various methods available for defining the thresholds and their main properties, see Fazio and Modica (2015) and Schluter (2021).

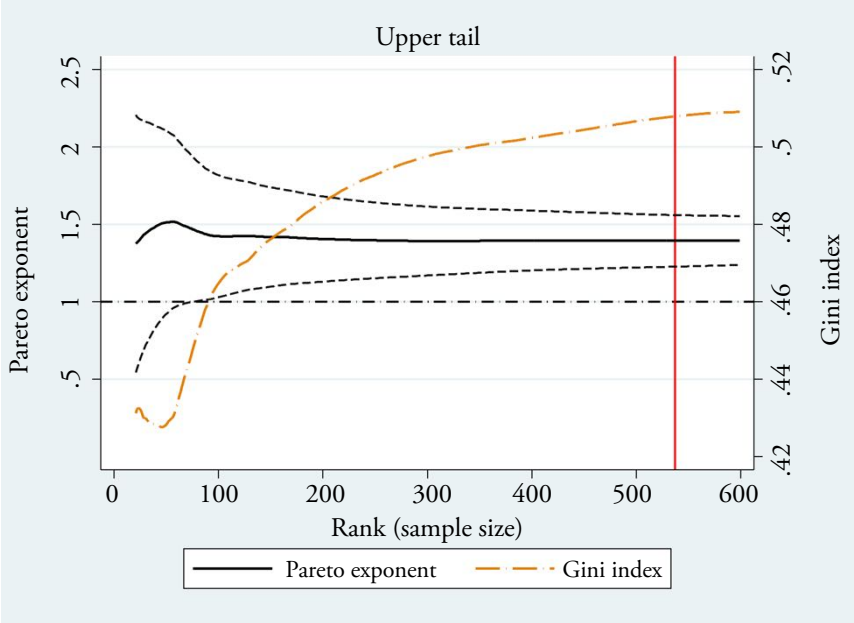
Schluter (2021); we also used the method of Ioannides and Skouras (2013) to estimate the threshold of the distribution that switches between a lognormal and a power-law distribution, and results hold[3]. Finally, the upper-tail graphs also include the calculation of the Gini index[4] by sample size—which does not impose a specific size distribution (Pareto for Zipf regressions)[5].

FIGURE 1.
Pareto exponent by sample size, US places in 2000 and 2010

a) US places in 2000 (whole sample)



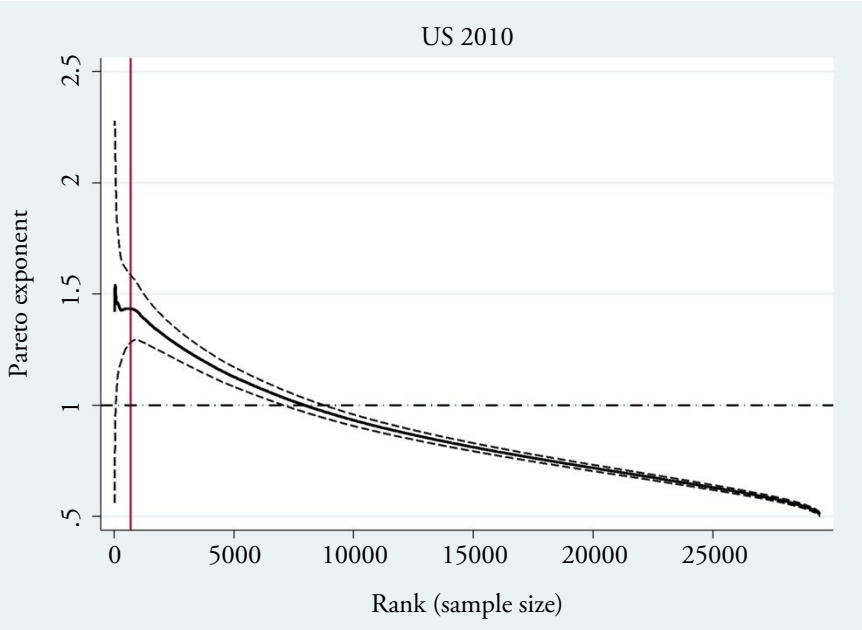
b) US places in 2000 (zoom)



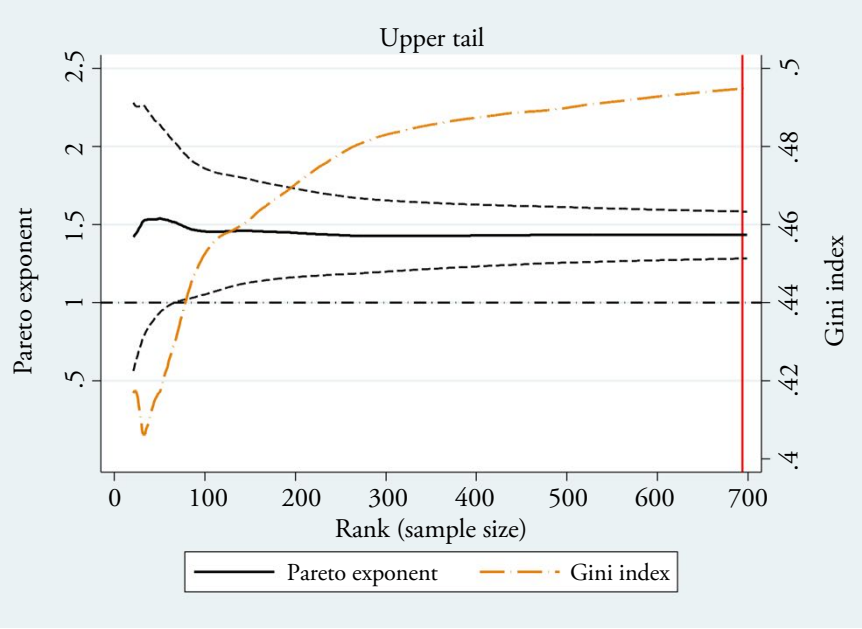
[3] These results are available from the authors upon request.
[4] The Gini coefficient is bounded by 0 (indicating perfect equality of city sizes) and 1 (indicating complete inequality).
[5] Results of the Clauset et al.'s (2009) test for a power law (not shown) reveal that the Pareto distribution provides a plausible fit to the data for the upper-tail city size distribution in all cases.

FIGURE 1. CONT.
Pareto exponent by sample size, US places in 2000 and 2010

c) US places in 2010 (whole sample)



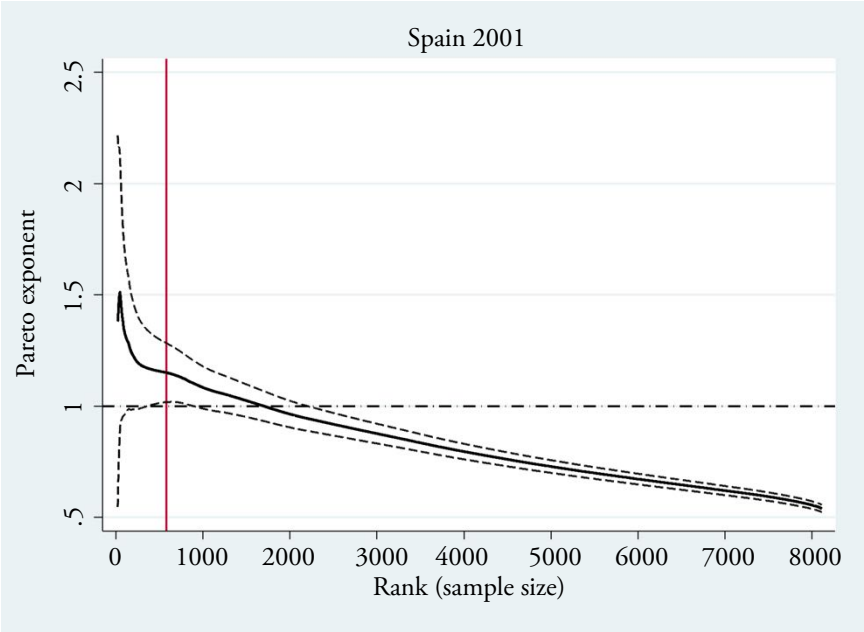
d) US places in 2010 (zoom)



Notes: The Pareto exponent is estimated utilising Gabaix and Ibragimov's Rank-1/2 estimator. Dashed lines represent the standard errors calculated by applying Gabaix and Ioannides's (2004) corrected standard errors: $\mathrm{GI\ s.e.} = \hat{\alpha}.(2/\mathrm{N})^{1/2}$, where *N* is the sample size. The vertical red lines indicate the threshold of the Pareto upper-tail, determined using Clauset et al. (2009)'s methodology. There are 537 places in the upper-tail in 2000 (the population threshold is 57,746) and 694 in 2010 (the population threshold is 55,156).

FIGURE 2.
Pareto exponent by sample size, Spanish municipalities in 2001 and 2011

a) Spanish municipalities in 2001 (whole sample)



b) Spanish municipalities in 2001 (zoom)

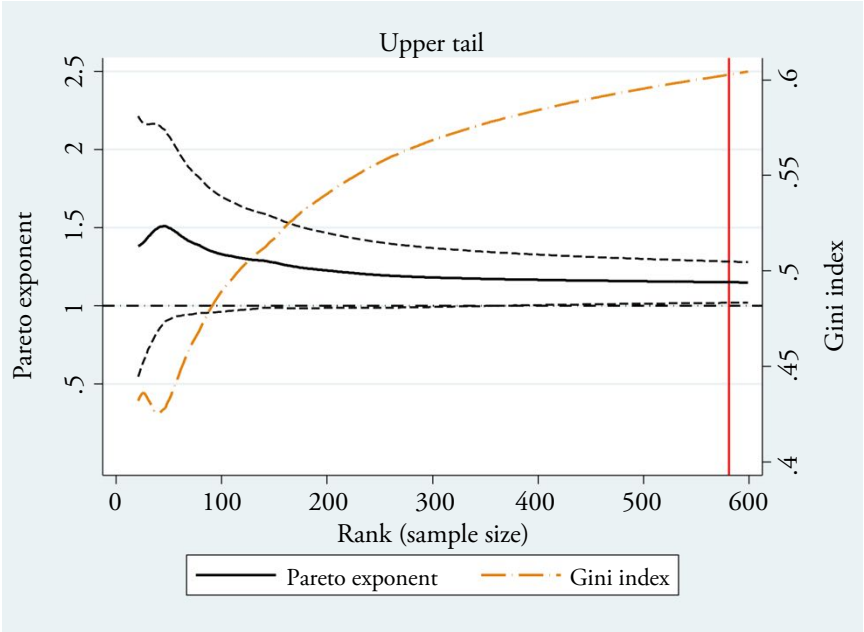
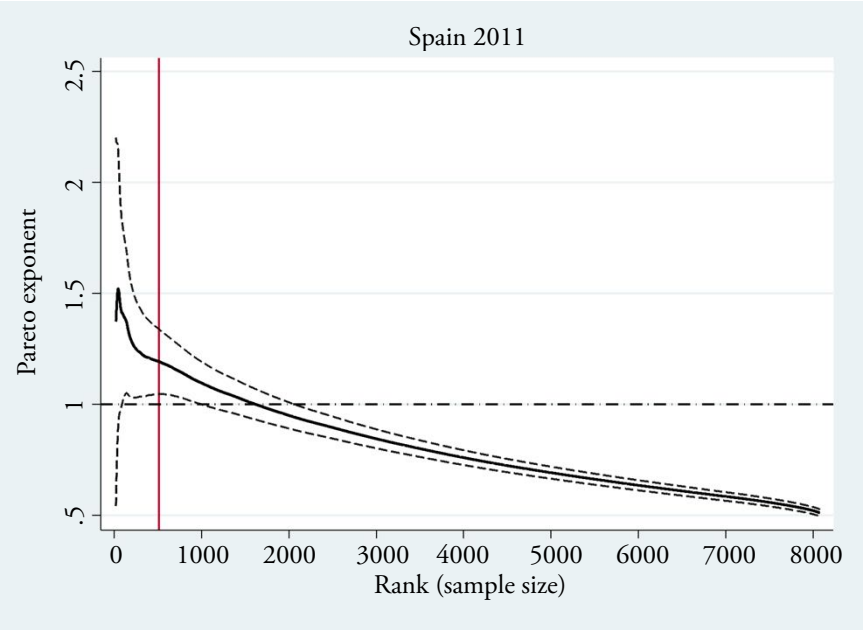
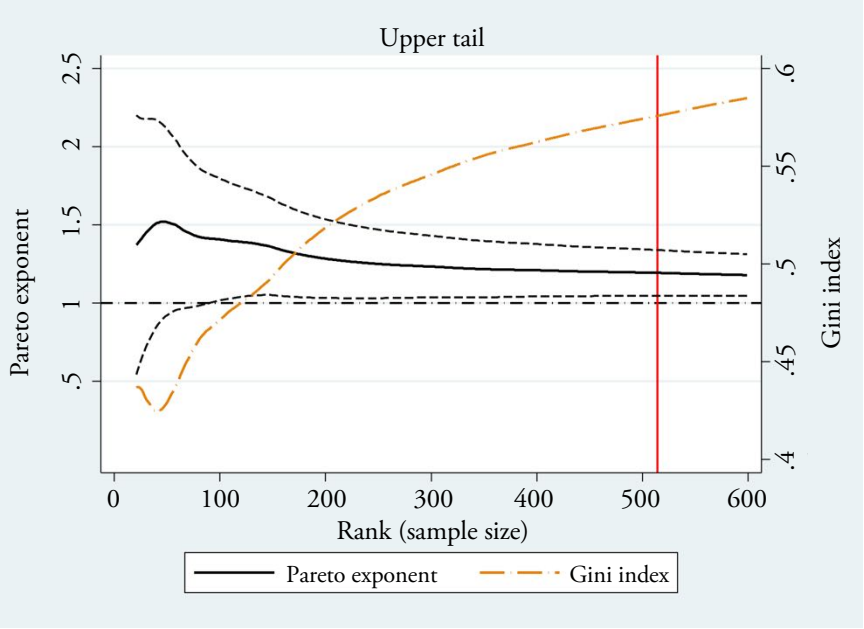


FIGURE 2. CONT.
Pareto exponent by sample size, Spanish municipalities in 2001 and 2011

c) Spanish municipalities in 2011 (whole sample)



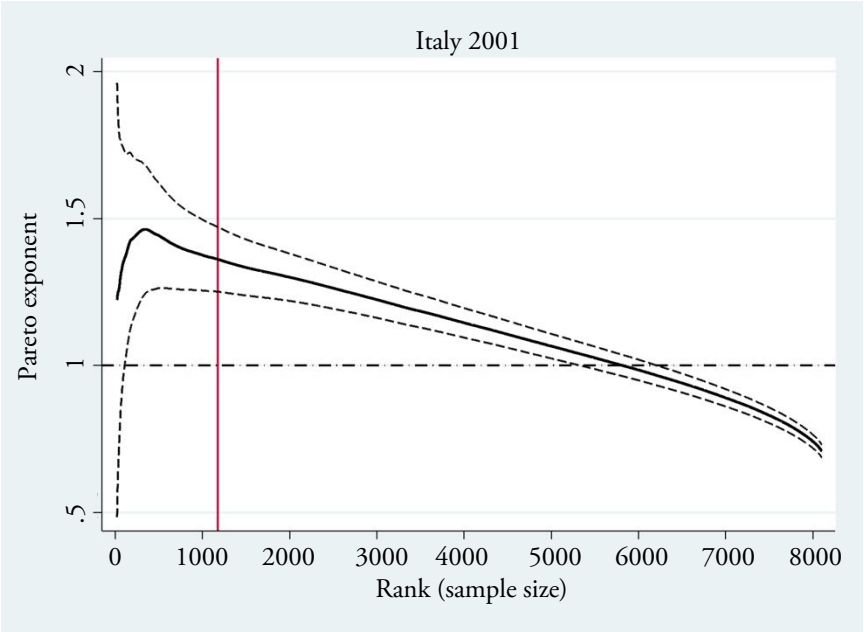
d) Spanish municipalities in 2011 (zoom)



Notes: The Pareto exponent is estimated utilising Gabaix and Ibragimov's Rank-1/2 estimator. Dashed lines represent the standard errors calculated applying by Gabaix and Ioannides's (2004) corrected standard errors: $\text{GI s.e.} = \hat{a}.(2/\text{N})^{1/2}$, where *N* is the sample size. The vertical red lines indicate the threshold of the Pareto upper-tail, determined using Clauset et al. (2009)'s methodology. There are 581 municipalities in the upper-tail in 2001 (the population threshold is 11,331) and 514 in 2011 (the population threshold is 15,583).

FIGURE 3.
Pareto exponent by sample size, Italian municipalities in 2001 and 2011

a) Italian municipalities in 2001 (whole sample)



b) Italian municipalities in 2001 (zoom)

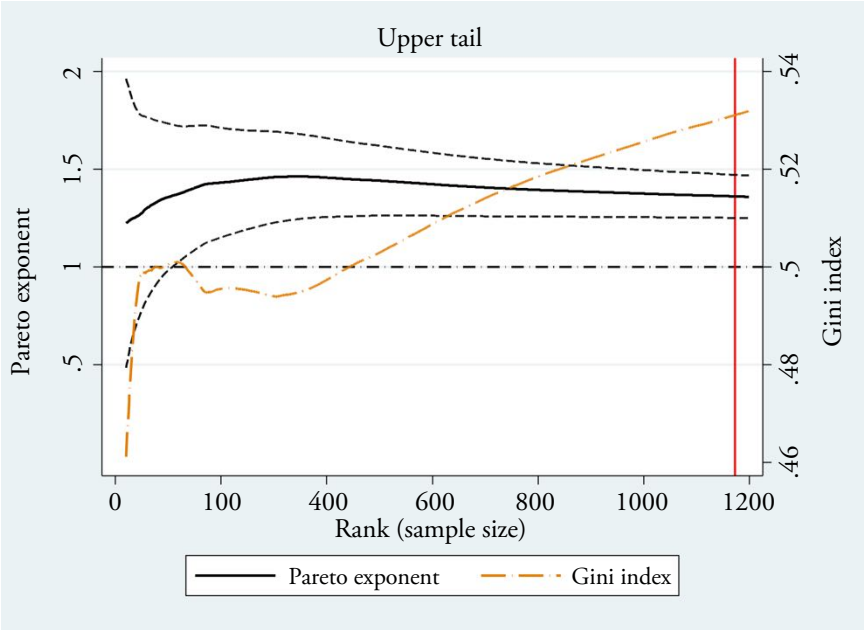
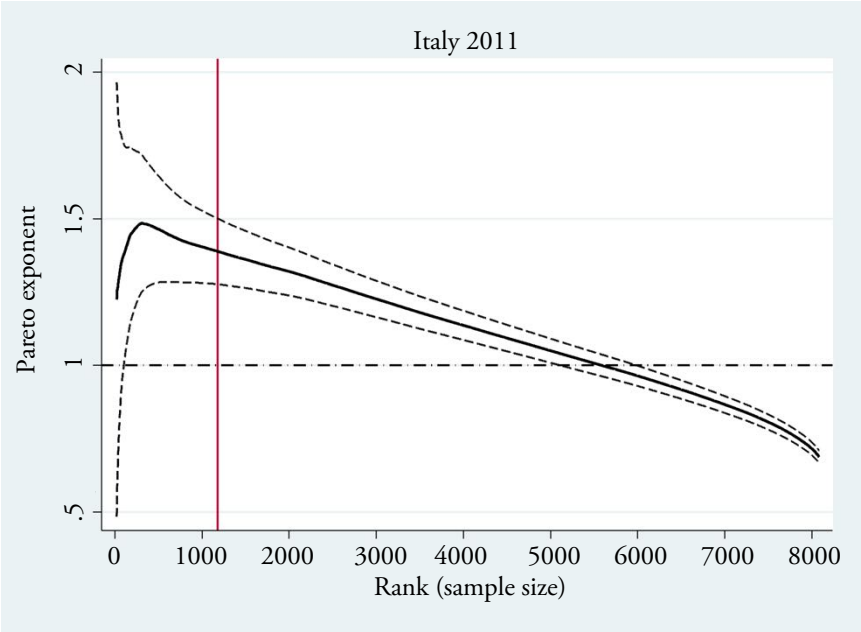
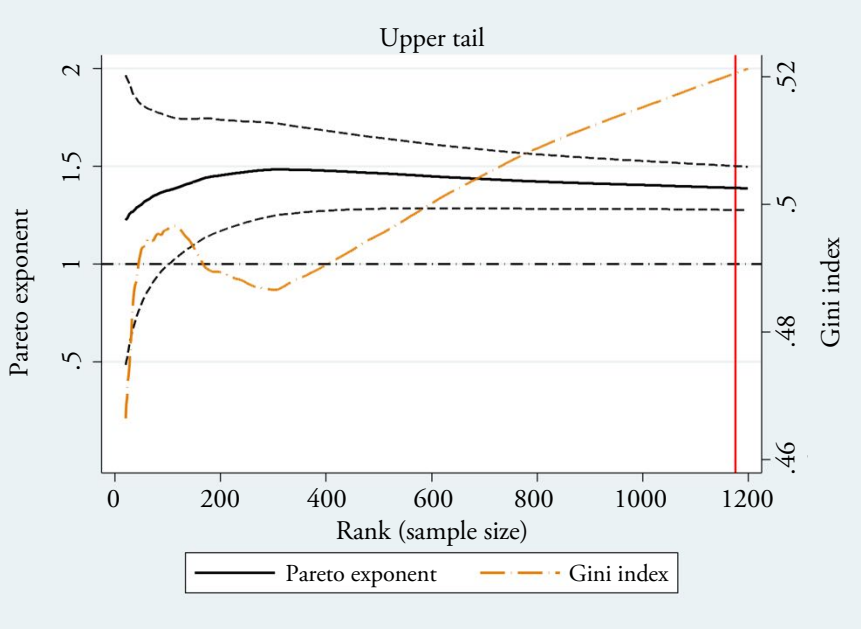


FIGURE 3. CONT.
Pareto exponent by sample size, Italian municipalities in 2001 and 2011

c) Italian municipalities in 2011 (whole sample)



d) Italian municipalities in 2011 (zoom)



Notes: The Pareto exponent is estimated utilising Gabaix and Ibragimov's Rank-1/2 estimator. Dashed lines represent the standard errors calculated by applying Gabaix and Ioannides's (2004) corrected standard errors: $\text{GI s. e.} = \hat{a}.(2/\text{N})^{1/2}$, where N is the sample size. The vertical red lines indicate the threshold of the Pareto upper-tail, determined using Clauset et al. (2009)'s methodology. There are 1,173 municipalities in the upper-tail in 2001 (the population threshold is 9,358) and 1,176 in 2011 (the population threshold is 10,282).

We can extract some general conclusions for the three countries. First, Zipf's law holds only for small sample sizes (approximately up to 100 cities) in all cases except for Spain in 2001, in which holds for all the upper-tail (Figure 2(b)), because the value of one only falls within the confidence bands for those small

sample sizes. Second, the estimation of the Pareto exponent for the upper-tail is independent of the sample size; in fact, it is not significantly different from a horizontal line. This finding implies that Proposition 1 in Eeckhout (2004) does not hold in the upper-tail in any case. Furthermore, this result also implies that the Pareto exponent does not provide any information about the degree of inequality in the upper-tail; the Pareto exponent is a flat line, which indicates that adding more cities does not increase or decrease inequality in city sizes. On the other hand, the Gini index clearly changes with the sample size. In most cases, the Gini index increases with increasing sample size (Figures 1(b), 1(d), 2(b) and 2(d)). That result points to increasing inequality with increasing sample size. However, the case of Italy is special: for sample sizes between 100 and 400 the Gini index decreases (Figures 3(b) and 3(d)). That finding implies growing homogeneity among city sizes; however, for sample sizes above the 400 largest cities, inequality increases like in the other countries.

Third, for sample sizes that extend beyond the upper-tail distribution (delineated by the vertical red line in the graphs) we observe exactly the same pattern in all cases: the estimate of the Pareto exponent decreases with sample size (Figures 1(a), 1(c), 2(a), 2(c), 3(a) and 3(c)). That is, once the sample size enters the lognormal body of the distribution and the lognormality assumption is fulfilled, Proposition 1 of Eeckhout (2004) is valid.

5. Conclusions

The current paradigm in the city size distribution literature states that, although most of the city size distribution is nonlinear, the Pareto distribution (and Zipf's law) holds for the largest cities (Giesen et al., 2010; Ioannides and Skouras, 2013). However, we found that Zipf's law holds only for small samples of the largest cities, not for the entire upper-tail.

Furthermore, we found that the Pareto exponent decreases with the sample size, but only for the lognormal body of the distribution. That finding partially supports Proposition 1 of Eeckhout (2004). For the upper-tail distribution, the exponent does not vary with the sample size. This result supports the current approach in the literature to estimating Zipf's law by considering only the upper-tail distribution, because if the upper-tail is Pareto-distributed the estimated exponent will be quite stable for all sample sizes within the upper-tail. However, a potential issue that remains is that, as the method for defining the Pareto upper-tail is still an open debate (Fazio and Modica, 2015; Schluter, 2021), different population thresholds (and sample sizes) across methods can lead to biased estimates of the Pareto exponent.

Acknowledgments

This research was funded by the Spanish Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación, MCIN/AEI/10.13039/501100011033 (projects PID2020-114354RA-I00 and PID2020-112773GB-I00), DGA (project S39_20R, ADETRE research group), and ERDF. Comments received from one anonymous referee have improved the version originally submitted. All remaining errors are ours.

References

Arshad, S., Hu, S., and Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and its Applications*, *492*, 75-92.

Arshad, S., Hu, S., and Ashraf, B. N. (2019). Zipf's law, the coherence of the urban system city size distribution: Evidence from Pakistan. *Physica A: Statistical Mechanics and its Applications*, *513*, 87-103.

Auerbach, F., (1913). Das Gesetz der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, *59*(74), 74-76.

Beirlant, J., P. Vynckier, and J. L. Teugels, (1996). Tail Index estimation, Pareto quantile plots, and regression diagnostics. *Journal of the American Statistical Association*, *91*(436), 1659-1667.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*, 661–703.

Clemente, J., González-Val, R., and Olloqui, I. (2011). Zipf's and Gibrat's laws for migrations. *The Annals of Regional Science*, *471*, 235-248.

Cottineau, C., (2017). MetaZipf. A dynamic meta-analysis of city size distributions. *PLoS ONE*, *8*, 1-22.

Cristelli, M., Batty, M., and Pietronero, L. (2012). There is more than a power law in Zipf. *Scientific Reports*, *2*, 1-7.

Düben, C., and Krause, M. (2021). Population, light, and the size distribution of cities. *Journal of Regional Science*, *61*(1), 189-211.

Eeckhout, J. (2004). Gibrat's Law for (All) Cities. *American Economic Review*, *94*(5), 1429-1451.

Fazio, G., and Modica, M. (2015). Pareto or log-normal? Best fit and truncation in the distribution of all cities. *Journal of Regional Science*, *55*(5), 736-756.

Fernholz, R. T., and Kramer, R. (2024). Racing to Zipf's law: Race and metropolitan population size 1910-2020. *Journal of Regional Science*, *64*(3), 649-670.

Gabaix, X., (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, *114*(3), 739-767.

Gabaix, X., and Ibragimov, R. (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics*, *29*(1), 24-39.

Gabaix, X., and Ioannides, Y. M. (2004). The evolution of city size distributions. In J. V. Henderson, and J. F. Thisse, (eds.), *Handbook of Regional and Urban Economics* (vol. 4, 2341-2378). Elsevier Science, North-Holland.

Giesen, K., Zimmermann, A., and Suedekum, J. (2010). The size distribution across all cities – double Pareto lognormal strikes. *Journal of Urban Economics*, *68*, 129-137.

González-Val, R., (2010). The evolution of US city size distribution from a long term perspective (1900-2000). *Journal of Regional Science*, *50*(5), 952-972.

González-Val, R., (2016). War size distribution: Empirical regularities behind conflicts. *Defence and Peace Economics*, 276, 838-853.

González-Val, R., Ximénez-de-Embún, D. P., and Sanz-Gracia, F. (2024). A long-term, regional-level analysis of Zipf's and Gibrat's laws in the United States. *Cities*, *149*, 104946.

Hackmann, A., and Klarl, T. (2020). The evolution of Zipf's Law for U.S. cities. *Papers in Regional Science*, *99*(3), 841-852.

Ioannides, Y. M., and Overman, H. G. (2003). Zipf's law for cities: an empirical examination. *Regional Science and Urban Economics*, *33*, 127-137.

Ioannides, Y. M., and Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, *73*, 18-29.

Luckstead, J., and Devadoss, S. (2014). Do the world's largest cities follow Zipf's and Gibrat's laws? *Economics Letters*, *125*(2), 182-186.

Nitsch, V., (2005). Zipf zipped. *Journal of Urban Economics*, *57*, 86-100.

Peng, G. (2010). Zipf's law for Chinese cities: Rolling sample regressions. *Physica A*, *389*, 3804-3813.

Peña, G., and Sanz-Gracia, F. (2021). Zipf's exponent and Zipf's law in the BRICS: A rolling sample regressions approach. *Economics Bulletin*, *41*, 2543-2549.

Puente-Ajovín, M., Ramos, A., and Sanz-Gracia, F. (2020). Is there a universal parametric city size distribution? Empirical evidence for 70 countries. *The Annals of Regional Science*, *65*, 727-741.

Schluter, C., (2021). On Zipf's law and the bias of Zipf regressions. *Empirical Economics*, *61*(2), 529-548.

Zanette, D. H., (2006). Zipf's law and the creation of musical context. *Musicae Scientiae*, *101*, 3-18.

Zipf, G. K., (1949). *Human behavior and the principle of least effort.* Addison-Wesley.

ORCID

Rafael González-Val https://orcid.org/0000-0002-2023-5726

Fernando Sanz-Gracia https://orcid.org/0000-0003-3725-0022

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 17-41
https://doi.org/10.38191/iirr-jorr.24.065

Articles

La dinámica de inversión regional en España: un estudio comparativo de las Comunidades Autónomas

Alberto Vaquero García, Santiago Lago Peñas**, María Cadaval Sampedro***, Patricio Sanchez Fernandez*****

Recibido: 14 de junio de 2024
Aceptado: 09 de octubre de 2024

Resumen:

Este trabajo tiene como objetivo realizar una descripción exhaustiva de la evolución de la inversión pública realizada por las Comunidades Autónomas en España entre 1984 y 2021, identificar patrones en sus políticas de inversión y evaluar la dinámica de sus fuentes de financiación: el ahorro, las transferencias y el déficit. Los resultados revelan diferencias significativas entre regiones, tanto en el nivel como en la dinámica de estas variables. Durante la Gran Recesión, se dio prioridad a los gastos sociales básicos, mientras que la recuperación económica mostró pautas diversas a nivel autonómico. Los fondos europeos y nacionales, como el Fondo de Compensación Interterritorial, supusieron un impacto positivo, revitalizando la inversión y el ahorro, convirtiéndose en un factor diferenciador en las Comunidades Autónomas.

Palabras clave: Inversión pública; transferencias de capital; endeudamiento; Comunidades Autónomas.
Clasificación JEL: H72; H73.

Regional Investment Dynamics: A Comparative Study of the Autonomous Communities in Spain

Abstract:

This paper aims to provide a comprehensive description of the evolution of public investment by the Autonomous Communities in Spain from 1984 to 2021, to identify patterns in investment policies and explain their underlying causes. Additionally, the paper evaluates the impact of savings, transfers, and deficits on public investment. The findings reveal distinct patterns in the level and dynamics of these variables across regions. Priority was given to essential welfare expenditures during the Great Recession, while the post-recession recovery displayed varied regional behaviors. The positive impact of European and national funds, such as the Interterritorial Compensation Fund, effectively revitalized investment and savings, which emerged as a key differentiating factor in chartered communities.

Keywords: Public investment; capital transfers; indebtedness; Autonomous Communities.
JEL Classification: H72; H73.

* Titular de Universidad. Departamento de Economía Aplicada. Facultad de Ciencias Empresariales y Turismo. Campus de Ourense. Universidade de Vigo. España. vaquero@uvigo.gal
** Catedrático de Universidad. Departamento de Economía Aplicada. Facultad de Ciencias Económicas y Empresariales. Campus de Santiago de Compostela. Universidade de Santiago de Compostela. España. santiago.lago.penas@usc.es
*** Contratada doctora. Departamento de Economía Aplicada. Facultad de Ciencias Económicas y Empresariales. Campus de Santiago de Compostela. Universidade de Santiago de Compostela. España. maria.cadaval@usc.es
**** Titular de Universidad. Departamento de Economía Financiera y Contabilidad. Facultad de Ciencias Empresariales y Turismo. Campus de Ourense. Universidade de Vigo. España. patricio@uvigo.gal
Autor para correspondencia: vaquero@uvigo.gal

1. Introducción

La política de inversión regional en España, implementada principalmente a través del Fondo de Compensación Interterritorial (FCI) y los Fondos europeos experimentó un impulso notable a principios de la década de 1990, coincidiendo con el proceso de descentralización de competencias autonómicas. Esto condujo a un aumento considerable de los recursos destinados a inversiones durante la construcción y consolidación del Estado autonómico, un crecimiento que se detuvo con la Gran Recesión. El año 2008 marcó un punto de inflexión en la inversión pública, se priorizó el gasto corriente sobre los gastos de capital, que no llegaron a recuperarse tras la crisis.

El principal objetivo de este artículo es ofrecer una caracterización exhaustiva de la evolución de la inversión pública realizada por las Comunidades Autónomas (CC. AA) durante el período 1984-2021. En particular analizamos las tendencias del gasto en inversión pública en las CC. AA y la dinámica de su financiación. Tras una revisión de la literatura y los determinantes que afectan las inversiones de los gobiernos subnacionales, presentamos una serie de hechos estilizados sobre la financiación de los gastos de capital. Adicionalmente, realizamos un análisis mediante el uso de una metodología de clúster que nos permite identificar patrones de comportamiento y un análisis σ-convergencia. En la sección de conclusiones se recogen los principales resultados y se proponen una serie de extensiones.

2. La inversión de capital de los gobiernos subcentrales

Desde la creación del Estado autonómico, se ha asignado una parte creciente del gasto público a las CC. AA, incluyendo inversiones en infraestructuras esenciales para su desarrollo y la gestión de funciones del Estado de Bienestar. Esta inversión se financia a través de tres vías, dos de ellas parcialmente endógenas (ahorro bruto y déficit), y una tercera que depende de otras administraciones públicas (transferencias de capital). Cuando los gastos corrientes de las CC. AA son inferiores a sus ingresos corrientes, se genera ahorro público, que puede destinarse a sufragar, total o parcialmente, los gastos de capital. En segundo lugar, las CC. AA pueden incurrir en déficit para financiar sus proyectos de inversión, especialmente cuando se prevé que estos generarán retornos futuros o cuando el monto de la inversión es significativo para ser cubierto en un solo ejercicio fiscal. En este contexto, las entidades subcentrales recurren a instrumentos de financiación que les permitan cubrir la inversión de capital mediante una combinación de ingresos actuales (corrientes) y futuros (deuda). La tercera vía de financiación son las transferencias de capital procedentes de otros niveles de gobierno. Unas transferencias que desempeñan un papel esencial para financiar la infraestructura, dado que las transferencias de nivelación y las condicionales ordinarias no suelen ser suficientes para cubrir el gasto requerido en infraestructuras (Martínez-Vázquez et al., 2022).

Una segunda cuestión importante a considerar son los fondos europeos. En esta investigación, que analiza la evolución de la inversión pública entre 1984 y 2021, es crucial tener en cuenta los cambios en la denominación y clasificación de las CC. AA que reciben estos fondos. Hasta 2006, se empleaba la terminología de "regiones Objetivo 1" para referirse a las CC. AA beneficiarias de los Fondos Europeos, que eran aquellas con PIB inferior al 75 % de la media comunitaria[1]. Desde 2007 y hasta 2013, la clasificación cambió a "regiones del objetivo convergencia"[2] para las que tenían un PIB inferior al 75% de la media comunitaria, y se introdujeron también las categorías de "competitividad"[3], "*phashing in*"[4] y "*phasing out*"[5]. A partir de 2014, la terminología se ajustó nuevamente, dividiéndose las regiones en "menos

[1] Las regiones Objetivo 1 eran Andalucía, Asturias, Canarias, Cantabria, Castilla y León, Castilla-La Mancha, Comunidad Valenciana, Extremadura, Galicia y Murcia.
[2] Las CC. AA beneficiarias de estos fondos eran Andalucía, Castilla-La Mancha, Extremadura y Galicia.
[3] Durante el período 2007-2013 fueron beneficiarias del Objetivo Competitividad las CC. AA de Aragón, Baleares, Cantabria, Cataluña, Comunidad de Madrid, Navarra, País Vasco y La Rioja.
[4] Las CC. AA beneficiarias de estos fondos eran aquellas que dejaron de pertenecer al Objetivo 1 por superar el 75% del PIB medio comunitario, esto es, Canarias, Castilla y León y Comunidad Valenciana.
[5] Este grupo estaba formado por aquellas CC. AA que se hubiesen situado por debajo del 75% de la media comunitaria a UE-15 y que han sobrepasado este porcentaje por haber disminuido la media después de la entrada de los 12 nuevos países. Se trata, por lo

desarrolladas", "más desarrolladas" y "de transición", cada una de ellas con criterios específicos para la asignación de fondos europeos. [6]

Finalmente, es precisa una referencia al sistema de financiación autonómica y, en particular, a la principal asimetría existente. El régimen foral, aplicable al País Vasco y Navarra, permite a estos territorios recaudar y administrar la práctica totalidad de los impuestos, realizando una transferencia al Estado en forma de cupo y de aportación, respectivamente. En contraste, el régimen común, que abarca a las otras quince CC. AA (con particularidades para Canarias), se basa en los tributos cedidos, compartidos y transferencias procedentes de otros niveles de gobierno. Las sucesivas reformas han fallado a la hora de considerar las variaciones en las necesidades y costes de infraestructura entre regiones, asumiendo erróneamente que la distribución del stock de capital público no requería ajustes. Esta omisión limita tanto la equidad como la eficiencia del sistema, dado que la evidencia empírica sugiere que la distribución actual del capital público no refleja adecuadamente las disparidades regionales en necesidades de inversión y costes (Herrero-alcalde et al., 2011).

3. Revisión de la literatura

En las últimas décadas, la literatura académica ha avanzado significativamente en la comprensión del papel de las transferencias de capital en las inversiones realizadas por los gobiernos, tanto nacionales como subnacionales. La efectividad de estas transferencias depende en gran medida de su diseño y de cómo las utilizan los gobiernos receptores. Bajo una doble categorización, condicionadas -que restringen el gasto a propósitos específicos- e incondicionadas -que permiten el gasto sin restricciones ni contrapartida-, la teoría del federalismo fiscal argumenta que estas últimas son adecuadas para afectar a la equidad vertical y horizontal, con plena autonomía para los receptores, mientras que las transferencias condicionadas son preferibles cuando se procura un objetivo específico y se espera que los gobiernos subnacionales refuercen políticas concretas. Así, tan importante es cómo utiliza cada gobierno los fondos que recibe como la definición clara de objetivos que se persiguen.

Tras las confirmaciones seminales de Solow (1956), Arrow (1962) o Romer (1986), numerosos trabajos avanzaron en el estudio del impacto que las transferencias de capital tienen sobre el desarrollo económico regional. Por otro lado, el artículo de Aschauer (1989) marcó el inicio de una línea de investigación que demuestra la importancia del capital público sobre el crecimiento económico, su impacto en la dinámica económica regional, la influencia en la productividad del sector privado y su eficacia como herramienta de política anticíclica. Barro (1991), Munnell (1992), Button (1998) y Peña (2008), entre otros, corroborraron este impacto positivo de las infraestructuras públicas en el desarrollo económico.

Otra línea de estudio complementaria aborda los criterios que determinan la asignación regional de las inversiones de capital con un enfoque espacial en los proyectos financiados con fondos europeos (De la Fuente, 2013, Murillo y Sosvilla 2005, Pastor et al., 2010, Mas-Ivars et al., 2010). El impacto territorial se suma a sus efectos sobre criterios de eficiencia y equidad en la distribución de bienes y servicios públicos (Holtz-Eakin y Rosen, 1989; Holtz-Eakin, 1994; Behrman y Craig, 1987), sin perder de vista las posibles motivaciones políticas que estudiaron Case (2001), Bosch y Espasa (1999) o Castells y Solé-Ollé (2005). También la inversión pública exhibe multiplicadores fiscales elevados en comparación con los derivados del consumo público (Alesina et al., 2015; Alloza, Burriel y Pérez, 2019; Ramey, 2020).

En investigaciones más recientes se destaca la importancia del capital público como elemento productivo determinante para las empresas privadas, capaz de desencadenar un ciclo virtuoso que contribuye a mitigar las tensiones económicas y aporta estabilidad en el corto y largo plazo (Asonuma y Joo, 2020; Bianchi et al., 2019). Incluso en contextos recesivos, la inversión pública puede generar efectos

tanto, de aquellas regiones que han dejado de pertenecer al objetivo convergencia por "efecto estadístico", esto es, Asturias, Murcia y las Ciudades Autónomas de Ceuta y Melilla.

[6] Extremadura se califica como región menos desarrollada; las regiones en transición son Andalucía, Canarias, Castilla-La Mancha, Murcia y la Ciudad Autónoma de Melilla, mientras que las regiones más desarrolladas son Aragón, Asturias, Baleares Cantabria, Cataluña, Castilla y León, Comunidad Valenciana, Comunidad de Madrid, Galicia, La Rioja, Navarra, País Vasco y la Ciudad Autónoma de Ceuta.

positivos en la demanda agregada, especialmente cuando el nivel de deuda no es excesivamente alto. Se respalda así el mecanismo keynesiano tradicional de estabilización (Noveli y Barcia, 2021; Gordon y Guerrón-Quintana, 2018).

El reconocimiento del impacto positivo de la inversión pública en la economía no ha evitado que las asignaciones presupuestarias destinadas a este fin se mantengan en niveles históricamente bajos en comparación con las tres últimas décadas (Abiad et al., 2015). Dos razones explican esta tendencia: el aumento previo del stock de capital público que justifica una menor inversión en el presente; en segundo lugar, y especialmente durante la última década, los procesos de consolidación presupuestaria derivados de la crisis financiera.

Aunque la primera de estas causas es de alcance general y afecta a todas las economías, la segunda ha tenido un impacto especialmente notorio en los países europeos (De Jong et al., 2017). Las medidas de consolidación fiscal implementadas han reducido significativamente las asignaciones destinadas a inversión, para preservar el gasto corriente. No obstante, esta estrategia plantea desafíos significativos para el desarrollo económico a corto plazo debido a la escasez de inversiones públicas.

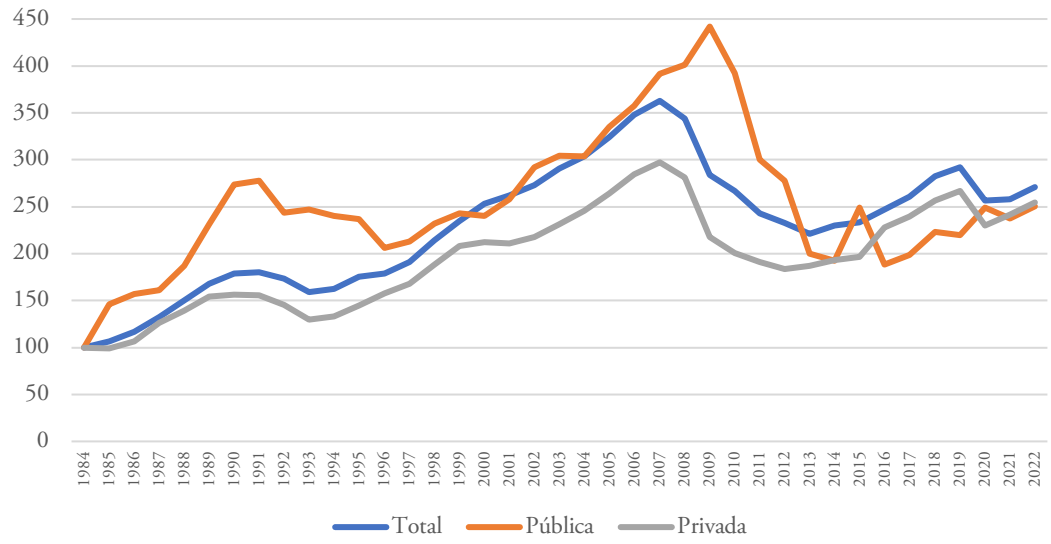
La ratio del gasto público sobre PIB ha experimentado un crecimiento significativo desde mediados del siglo XX en los países que conforman la UE-15. Entre 1960 y 2021, pasaron del 23 % al 50,6 %. La inversión pública en la Unión Europea muestra un comportamiento distinto en los últimos años, y España no ha sido ajena a esta tendencia[7]. Mientras que la UE-15 disminuyó su inversión pública respecto al PIB del 10 % en 2012 al 8,5 % en 2019, España experimentó una reducción aún mayor, pasando durante estos dos años del 13,2 % al 6,7 %. Esta cifra, que incluye las transferencias de capital, se situó por debajo del percentil 25 en la distribución europea, con un peso relativo del 2,8 % en 2019, un punto por debajo del promedio de la UE-15 y 1,5 puntos inferior a la UE-27, superando solo a Portugal en esta comparativa. Una década atrás, la situación era muy diferente: la inversión pública en España, tanto en términos absolutos como en relación con el PIB, se encontraba en el percentil 75 de los países europeos. La dinámica posterior a la crisis financiera resultó en una disminución de la inversión pública por debajo del nivel de depreciación de los activos, lo que en la práctica ha supuesto una reducción del stock de capital neto (Mas-Ivars y Pérez García, 2021). Esta reducción se ha reflejado de manera más acentuada, especialmente en algunas CC. AA -por ejemplo, la Comunidad Valenciana, Cataluña o la Comunidad de Madrid- donde se ha producido un marcado retroceso en el gasto en inversión hasta niveles no observados desde hace más de dos décadas.

Tras esta constatación, es relevante realizar una comparación temporal entre la inversión pública y privada. La Figura 1 ilustra la evolución de ambas variables desde 1984 hasta 2022. La inversión pública experimentó un aumento significativo (Mas-Ivars et al., 2015), coincidiendo con un período de fuerte expansión económica. A partir de entonces, comenzó a disminuir, hasta alcanzar su punto más bajo en 1996. Por su parte, la inversión privada, aunque siempre inferior a la pública hasta el año 2000, muestra una tendencia al alza. Entre el año 2000 y el 2006, los niveles de inversión pública y privada se equipararon, registrando ambas un comportamiento creciente. En 2007, se observa el inicio de un desacoplamiento entre los niveles de inversión pública y privada. Mientras que la segunda sufrió una notable disminución debido a las medidas contracíclicas adoptadas, la primera continuó aumentando hasta 2009, como indican Mas-Ivars y Pérez García (2021). A partir de 2009, la inversión pública comenzó a declinar significativamente, reduciéndose a menos de un tercio del nivel registrado en 2009 para 2013. En ese mismo año, el crecimiento acumulado de la inversión privada superó al de la pública. Este patrón se mantuvo hasta 2020, con dos excepciones en 2015 y 2019. En 2021, último año del período analizado, la inversión pública se equiparó a la privada. En términos del PIB, entre 2014 y principios de 2020, se observó

[7] España, a partir de los años 80, inició un proceso de convergencia que se intensificó en los años 90. En 2021, su peso superaba al del Reino Unido, Dinamarca y Suecia, entre otros, aunque este dato estaba aún afectado por un PIB reducido que no había recuperado su valor prepandemia (en 2023 estaba tres puntos porcentuales por debajo de la media de la UE27). Con todo, si bien el proceso de convergencia español ha superado en 12 puntos la media europea en términos nominales, no ha sido así en términos per cápita reales, donde la brecha ha aumentado desde 1999 como consecuencia de las dinámicas demográficas y de renta (Pérez y Solera, 2017). Sin embargo, entre 2012-2019 el flujo de inversión pública no fue suficiente para compensar la reducción del consumo del capital fijo de la administración pública, lo que explica la pérdida del stock relativo con respecto al PIB a partir de 2014.

una recuperación significativa, que fue seguida por importantes descensos a partir del primer trimestre de 2020 como consecuencia de la pandemia del Covid-19.

FIGURA 1.
Inversión bruta pública y privada real. España (1984-2022) (1984=100)



Fuente: Mas-Ivars y Pérez (2022).

4. LA INVERSIÓN AUTONÓMICA Y SU FINANCIACIÓN: HECHOS ESTILIZADOS ENTRE 1984-2021

La relación entre la inversión autonómica y su financiación es una cuestión tratada en algunos trabajos previos (Conde Ruiz et al., 2016; Fernández Leiceaga y Lago Peñas, 2013; Lago Peñas y Vaquero García, 2016; Lago Peñas et al., 2017a y 2017b). El punto de partida para elaborar esta sección es la construcción de una base de datos que incluye información anual sobre la evolución del gasto en operaciones de capital, ahorro público, déficit presupuestario y transferencias de capital en las CC.AA. españolas desde 1984 hasta 2021, a partir de la información que proporciona el Ministerio de Hacienda. Los datos están ajustados a euros constantes de 2015 utilizando el deflactor del PIB y se presentan en términos per cápita para facilitar comparaciones, usando la información sobre población que ofrece el Instituto Nacional de Estadística (INE) (Ver Cuadro 1). El estudio divide el período en dos fases clave: 1984-2007 y 2008-2021.

El cuadro 2 muestra una correlación positiva significativa entre el gasto de capital neto[8] y el ahorro público, así como una correlación negativa entre el ahorro y el déficit. Esta interacción destaca la relación directa entre la inversión y la capacidad de ahorro de las CC.AA., y la influencia negativa del déficit en los niveles de ahorro. Además, se aprecia una correlación positiva, aunque moderada, entre el gasto de capital neto de los gobiernos regionales y las subvenciones de capital. El cuadro 4 del anexo presenta la distribución temporal de los Fondos Europeos entre las CC. AA. y del FCI, información clave para entender la evolución y distribución de las transferencias de capital a lo largo del tiempo (Lago Peñas et al., 2015).

[8] Es el sumatorio de los capítulos 6 y 7 del presupuesto de gastos.

CUADRO 1.
Variables utilizadas y fuente de información

Acrónimo	Variable	Unidades	Definición	Fuente
P	Población	Habitantes	Población residente	Instituto Nacional de Estadística (2024)
GK	Gastos de capital neto de los gobiernos regionales	€/habitante	Gasto de capital neto de los gobiernos regionales	Ministerio de Hacienda (2024)
S	Ahorro público	€/habitante	Diferencia entre ingresos y gastos corrientes de las CCAA	
DEF	Déficit público	€/habitante	Desequilibrio fiscal total de las administraciones regionales (los valores negativos corresponden a superávit)	
TK	Subvenciones/Transferencias de capital	€/habitante	Subvenciones de capital recibidas por los gobiernos regionales	

Fuente: Elaboración propia.

CUADRO 2.
Matriz de correlaciones

	GK	TK	S	DEF
GK	1	0,420	0,644	0,024
TK	0,420	1	0,106	-0,117
S	0,644	0,106	1	-0,694
DEF	0,024	-0,117	-0,694	1

Fuente: Elaboración propia.

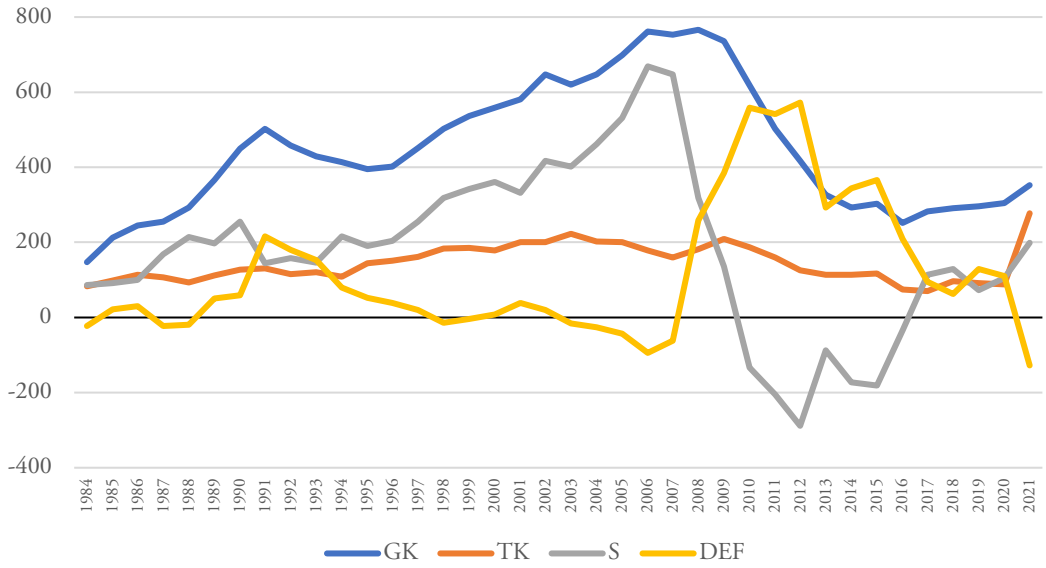
El impacto de la Gran Recesión es determinante para comprender la evolución de las finanzas autonómicas. A partir de 2007, se observa un retroceso en el ahorro y un aumento del déficit público o necesidad de financiación [9] (ver figura 2) debido a la disminución de los ingresos y la rigidez de los compromisos de gasto, especialmente en servicios fundamentales como la sanidad, la educación y los servicios sociales. Esta rigidez, junto con la caída de ingresos por la crisis económica, provocó una necesidad creciente de financiación que condujo a déficits más elevados.

En el período 1984-2007, Navarra lideró la inversión en capital per cápita con un promedio de 1.047 euros por habitante, que se justifica por sus competencias específicas, como la gestión de la Red de carreteras de Interés General del Estado (RIGE), y su mayor capacidad financiera. Otras comunidades, como Galicia, Extremadura y Asturias, también presentaron niveles elevados de inversión, beneficiándose del FCI y de Fondos Europeos. En el extremo opuesto, Madrid y Cataluña, que recibieron menos recursos de estos fondos, presentaron niveles de inversión considerablemente más bajos. En el período 2008-2021, la inversión en capital disminuyó en la mayoría de las CC.AA., con un marcado descenso en Cataluña y la

[9] El déficit público o necesidad de financiación se define como la diferencia entre los ingresos y los gastos no financieros.

Comunidad Valenciana, regiones que enfrentaron graves déficits y priorizaron el gasto corriente. Navarra siguió destacándose por su alta inversión en capital (687 euros por habitante), mientras que Madrid y Cataluña se ubicaron entre las que menos invirtieron. Respecto a las transferencias de capital, aunque inicialmente presentaron una mayor estabilidad en comparación con otras variables, experimentaron un cambio significativo que alteró su tendencia positiva, convirtiéndola en negativa. Este cambio provocó una notable disminución en el gasto en inversión de capital, afectado por los recortes impuestos por el marco de estabilidad presupuestaria, reduciendo dicho gasto a casi la mitad.

FIGURA 2.
Evolución de las medias transversales de los gastos de capital, las subvenciones de capital, el ahorro público y el déficit presupuestario entre 1984-2021. Cifras en euros constantes per cápita



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

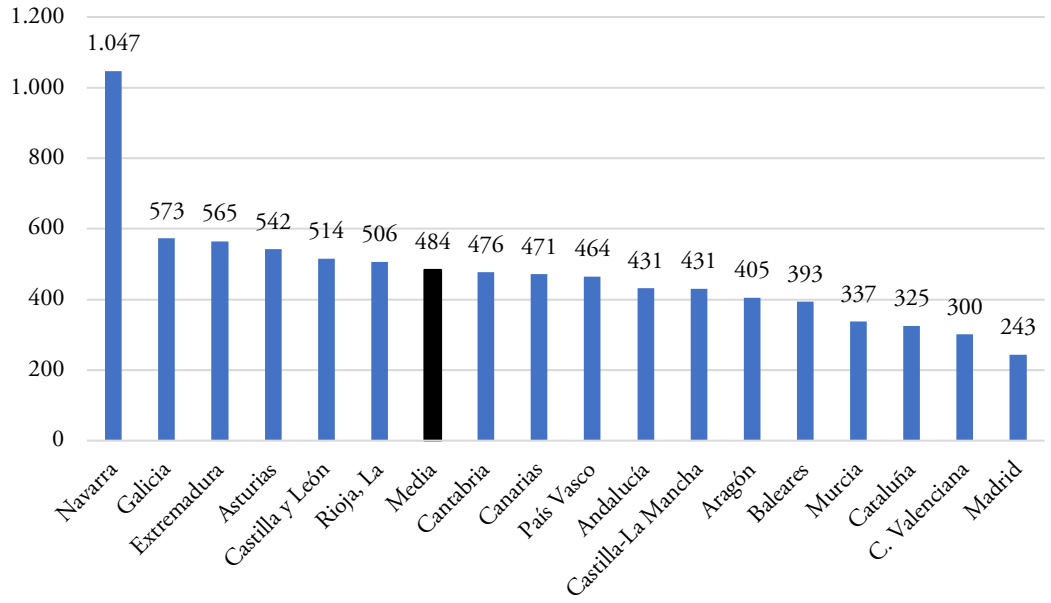
Entre 1984 y 2007, todas las CC.AA. registraron ahorros positivos, con Navarra y el País Vasco a la cabeza (Figura 3). Sin embargo, entre 2008 y 2021, muchas comunidades experimentaron ahorros negativos, particularmente en Cataluña, Comunidad Valenciana y Murcia, que se vieron especialmente afectadas por la crisis financiera (Figura 4). El análisis muestra que, además de las forales, las CC.AA. que han sido beneficiarias de fondos europeos han logrado mantener niveles más estables de ahorro e inversión en comparación con aquellas que no han sido elegibles para estos recursos.

Finalmente, hay que referirse al déficit público, indicador clave en las finanzas autonómicas, y cuyo comportamiento ha sido objeto de un extenso análisis. Según De la Fuente (2013), entre 2003 y 2007, los ingresos de CC.AA. crecieron a la par que los gastos. Sin embargo, a partir de 2007, el déficit respondió al desacople entre la rigidez de los compromisos de gasto y la evolución negativa de los ingresos (Ruiz-Huerta y García, 2012). Este desajuste se debe, en gran medida, al elevado peso de las competencias de las CC.AA. en áreas fundamentales como sanidad, educación y servicios sociales, que constituyen aproximadamente dos tercios de sus presupuestos.

A pesar de la implementación de la normativa de estabilidad presupuestaria y la reforma del artículo 135 de la Constitución en 2011, no se logró contener las desviaciones fiscales, debido, en gran parte, a la flexibilidad normativa y a la amplitud en su interpretación (Díaz Roldán y Cuenca, 2014; Lago Peñas, 2015). La recuperación económica que precedió a la pandemia tampoco permitió un control efectivo del déficit autonómico. Algo que, en buena medida, se explica porque desde 2012 se implementaron

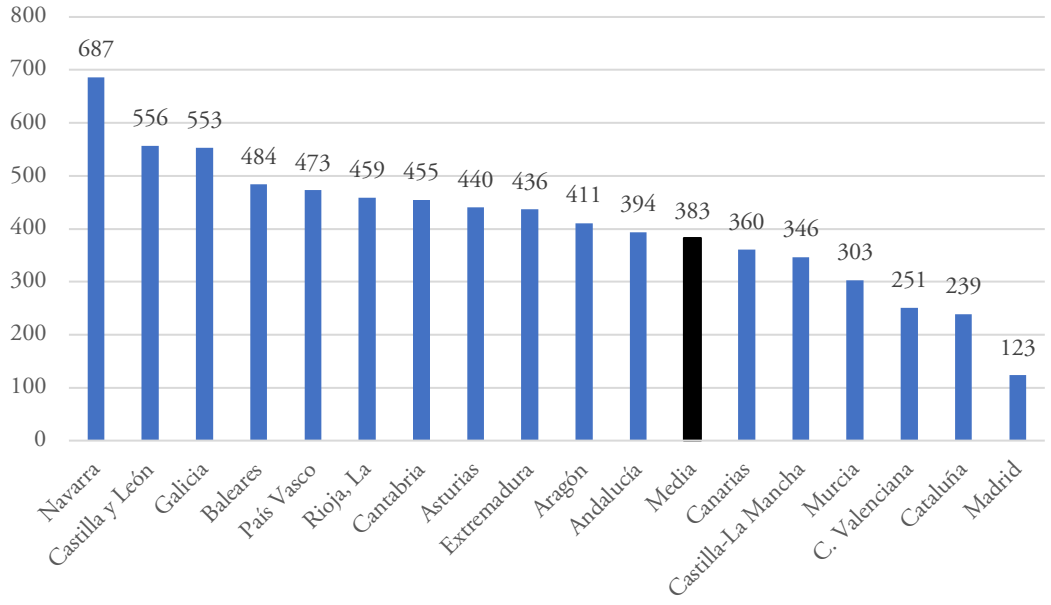
mecanismos de asistencia financiera a través de fondos extraordinarios, que han alimentado una cierta laxitud en la restricción presupuestaria de las CC. AA.

FIGURA 3.
Media anual del gasto en capital por CC. AA (1984-2007), euros constantes por habitante, año base 2015



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

FIGURA 4.
Media anual del gasto en capital CC. AA (2008-2021), euros constantes por habitante, año base 2015



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

En resumen, el análisis de los resultados correspondientes al período 1984-2021 revela una evolución desigual del déficit y las transferencias de capital entre las CC. AA. Durante el primer período, 1984-2007, algunas comunidades como el País Vasco y Navarra presentaron ligeros superávits. En contraste, desde 2008 a 2021, todas las comunidades experimentaron déficits, siendo especialmente significativos los casos de Baleares, Cataluña, Murcia y la Comunidad Valenciana. Las transferencias de capital jugaron un papel crucial en comunidades como Extremadura, Galicia y Asturias, mientras que otras, como Madrid y Valencia, recibieron asignaciones más bajas. Además, durante el segundo período, se observó una disminución en la relevancia de las transferencias de capital, a excepción de las regiones tradicionalmente beneficiarias de fondos europeos (ver Cuadro 3).

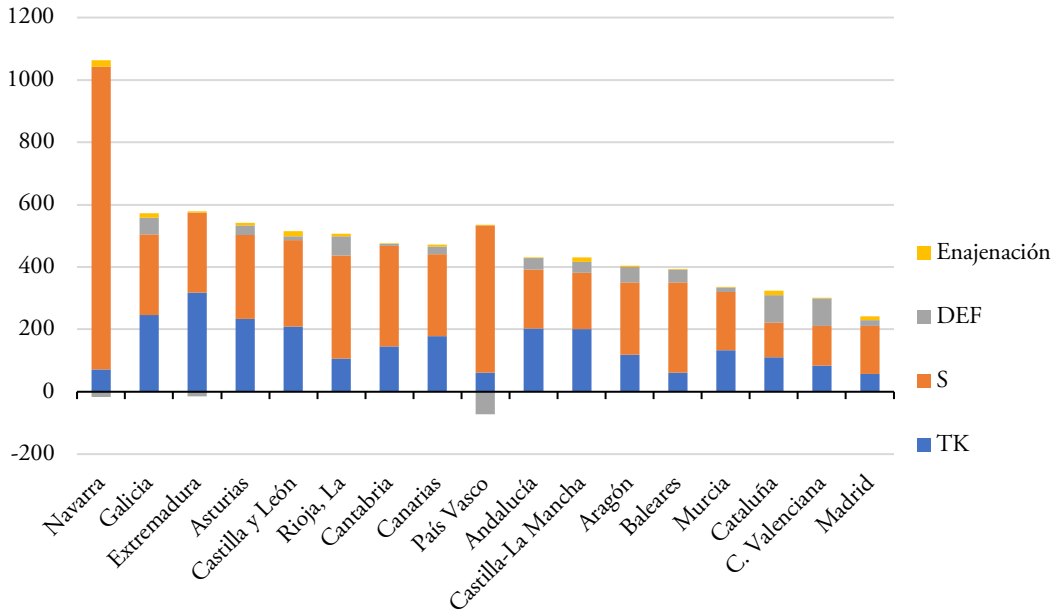
CUADRO 3.
Media anual por CC. AA de los gastos de capital, ahorro, déficit, y transferencias de capital (1984-2007 y 2008-2021), euros per cápita constantes, año base 2015

	Gastos de capital		Ahorro		Déficit		Transferencias de capital	
	1984 - 2007	2008 - 2021	1984 - 2007	2008 - 2021	1984 - 2007	2008 - 2021	1984 - 2007	2008 - 2021
Andalucía	431	394	189	21	36	187	203	181
Aragón	405	411	231	-79	49	338	120	150
Asturias	542	440	269	58	29	151	235	229
Baleares	393	484	288	11	42	402	62	68
Canarias	471	360	264	66	24	116	177	177
Cantabria	476	455	323	79	7	294	145	80
Castilla y León	514	556	276	96	13	272	210	175
Castilla-La Mancha	431	346	181	-235	35	397	200	177
Cataluña	325	239	110	-269	88	413	111	89
C. Valenciana	300	251	128	-276	88	477	83	49
Extremadura	565	436	257	-131	-15	264	318	299
Galicia	573	553	259	179	54	146	246	220
Madrid	243	123	154	-6	18	98	57	21
Murcia	337	303	188	-236	14	425	133	108
Navarra	1.047	687	970	436	-17	164	72	82
País Vasco	464	473	473	128	-72	252	61	89
Rioja, La	506	459	331	125	61	213	106	120

Fuente: Elaboración propia con datos del Ministerio de Hacienda.

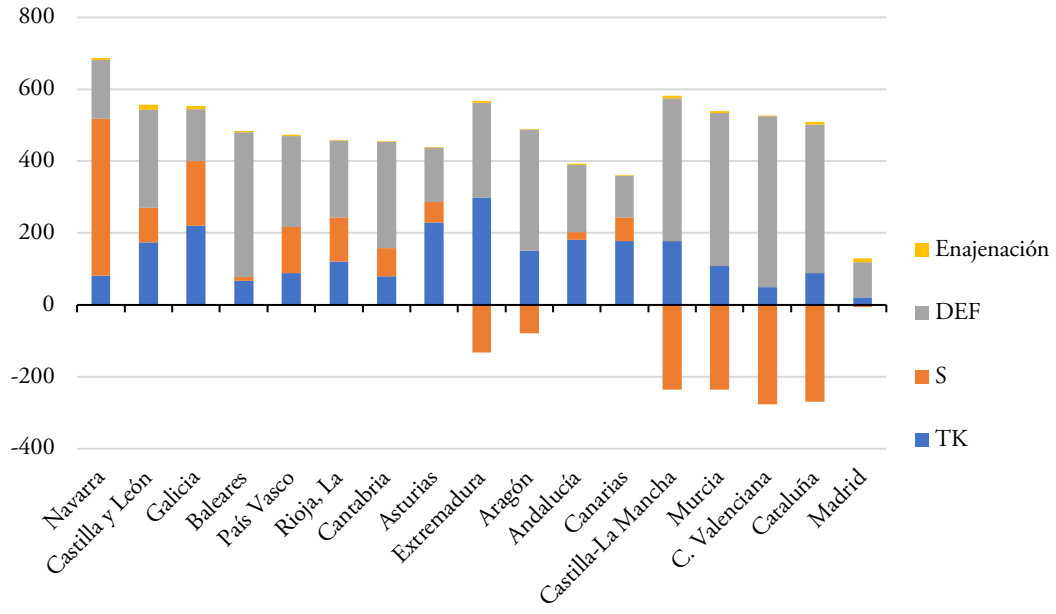
Las Figuras 5 y 6 sintetizan la dinámica de las diversas fuentes de financiación en cada CC.AA. para el período 1984-2007 y 2008-2021, respectivamente. Durante el primer período, el ahorro fue una fuente clave de financiación, especialmente en Navarra. En cambio, en el segundo período, varias CC.AA., como Castilla-La Mancha, Murcia, Comunidad Valenciana y Cataluña, presentaron ahorro negativo. Las transferencias de capital constituyeron la segunda fuente más importante de financiación de los gastos de capital, desempeñando un papel fundamental en el primer período. No obstante, su relevancia se redujo considerablemente durante el segundo, afectada por la crisis financiera y la respuesta a la pandemia. En cuanto a los saldos presupuestarios, se observa una transición de superávits moderados a déficits generalizados en la mayoría de las regiones. La contribución de las enajenaciones reales fue marginal en ambos períodos.

FIGURA 5.
Financiación de los gastos de capital: Transferencias de capital, ahorro, déficit y enajenación de inversiones. Promedio para el período 1984-2007, euros per cápita constantes, año base 2015



Nota: Enajenación: Enajenación de inversiones reales; DEF: Déficit, S: Ahorro; TK: Transferencias/Subvenciones de capital.
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

FIGURA 6.
Financiación de los gastos de capital: Transferencias de capital, ahorro, déficit y enajenación de inversiones. Promedio para el período 2008-2021, euros per cápita constantes, año base 2015

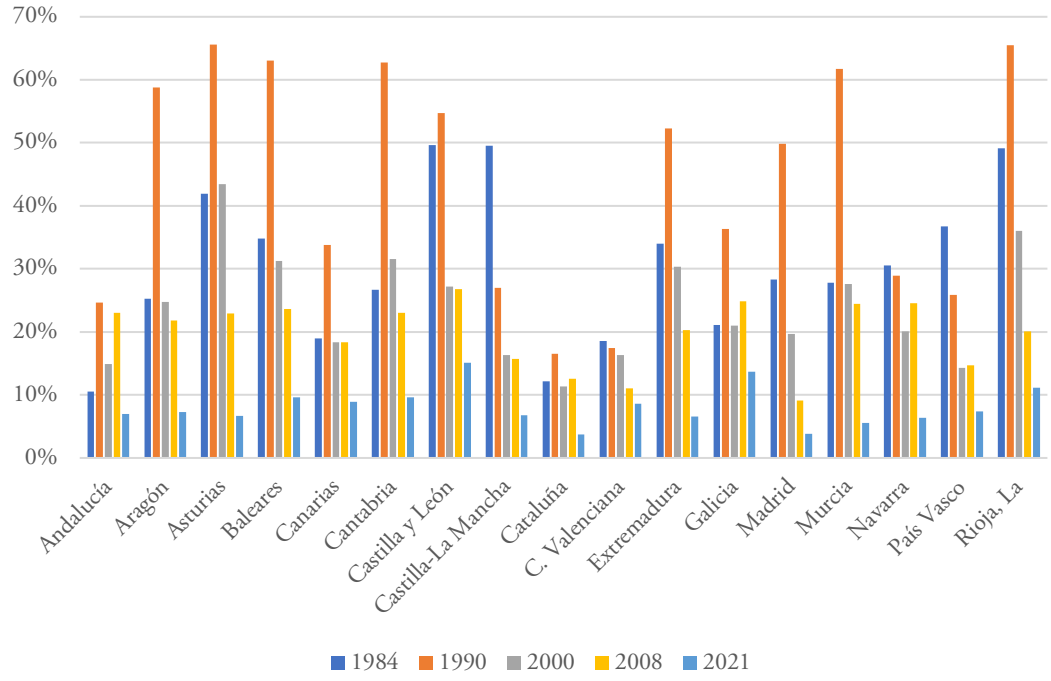


Nota: Enajenación: Enajenación de inversiones reales; DEF: Déficit, S: Ahorro; TK: Transferencias/Subvenciones de capital.
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

Estos resultados revelan un patrón de financiación de los gastos de capital que reflejan un comportamiento distintivo. El caso de Navarra resalta por su capacidad de ahorro sostenido, ya que nunca fue considerada región beneficiaria de Fondos Europeos y solo recibió financiación del FCI durante el quinquenio 1984-1989. En contraste, otras regiones como Andalucía, Galicia, Castilla y León, Extremadura, Asturias, Canarias, o Castilla-La Mancha, al haber sido beneficiarias de los Fondos Europeos, dependieron en mayor medida de las transferencias de capital.

Para evaluar el peso relativo de las cuatro variables a lo largo del tiempo, se han relativizado los importes usando los ingresos no financieros en varios años (1984, 1990, 2000, 2008 y 2021). La Figura 7 ilustra que, hasta 1990, la proporción de gastos de capital respecto a los ingresos no financieros aumentó, con incrementos significativos en comunidades como Aragón, Asturias, Baleares, Cantabria, Extremadura, Comunidad de Madrid, Murcia y La Rioja. Sin embargo, en el año 2000 se registró una marcada disminución en esta relación para todas las CC.AA. Desde entonces, la asignación de gastos de capital en términos reales ha continuado decreciendo, especialmente en aquellas regiones que habían experimentado los mayores incrementos hasta 1990.

FIGURA 7.
Gastos de capital sobre ingresos no financieros en 1984, 1990, 2000, 2008 y 2021. Datos en porcentaje



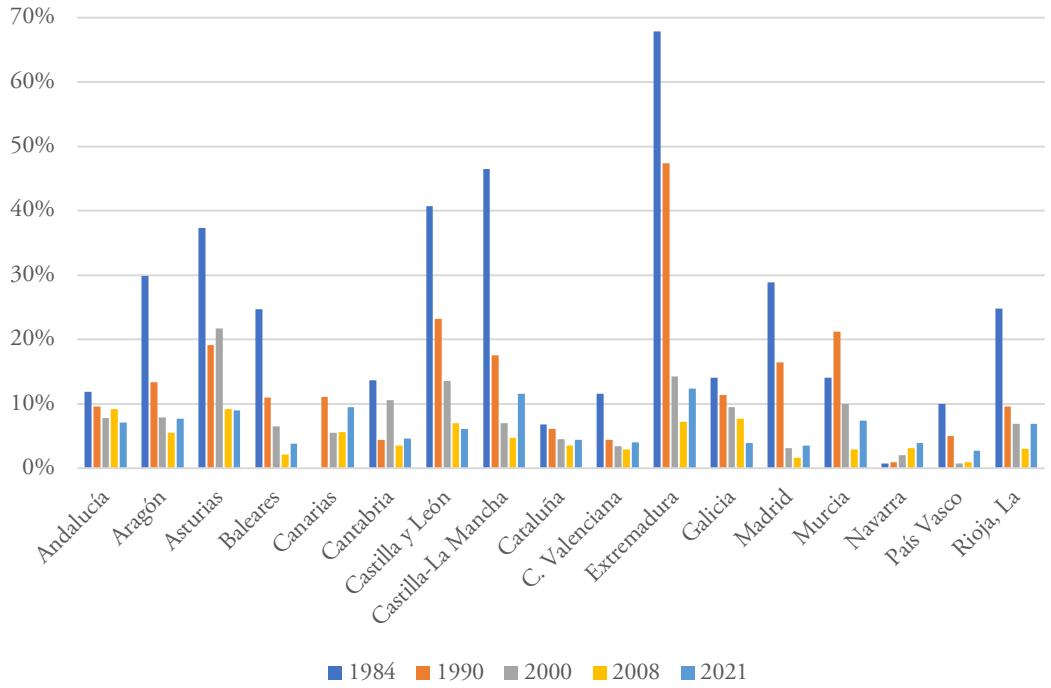
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

La Figura 8 revela una tendencia decreciente a lo largo del período analizado. Aunque esta disminución es generalizada, se destaca especialmente en Extremadura, Castilla-La Mancha, Castilla y León y Asturias. Por otro lado, en Andalucía, Canarias, Cantabria, Cataluña, Comunidad Valenciana, Galicia y el País Vasco, la reducción ha sido menos pronunciada, siendo Navarra la única comunidad que registra un incremento en esta ratio.

En lo que atañe al déficit, la Figura 9 refleja que la mayoría de las CC.AA. mostraron superávit en 1984, transformándose en déficits en 1990, particularmente acentuados en Cantabria, Comunidad de Madrid, Murcia y La Rioja. Desde el año 2000, este fenómeno se ha moderado, aunque persisten déficits, siendo menos pronunciados. Cabe destacar que Cataluña, Murcia y la Comunidad Valenciana han mostrado déficits relativos durante todos los años estudiados, incluso en 2021, cuando el resto de las CC.AA. presentaron superávits.

Finalmente, se centrará en la dinámica de la ratio del ahorro en la Figura 10. Se mantiene positiva en la mayoría de las CC.AA., con excepciones en Baleares, Cataluña, Comunidad Valenciana, Extremadura y Murcia.

FIGURA 8.
Transferencias de capital sobre ingresos no financieros,1984-1990-2000-2008-2021. Datos en porcentaje

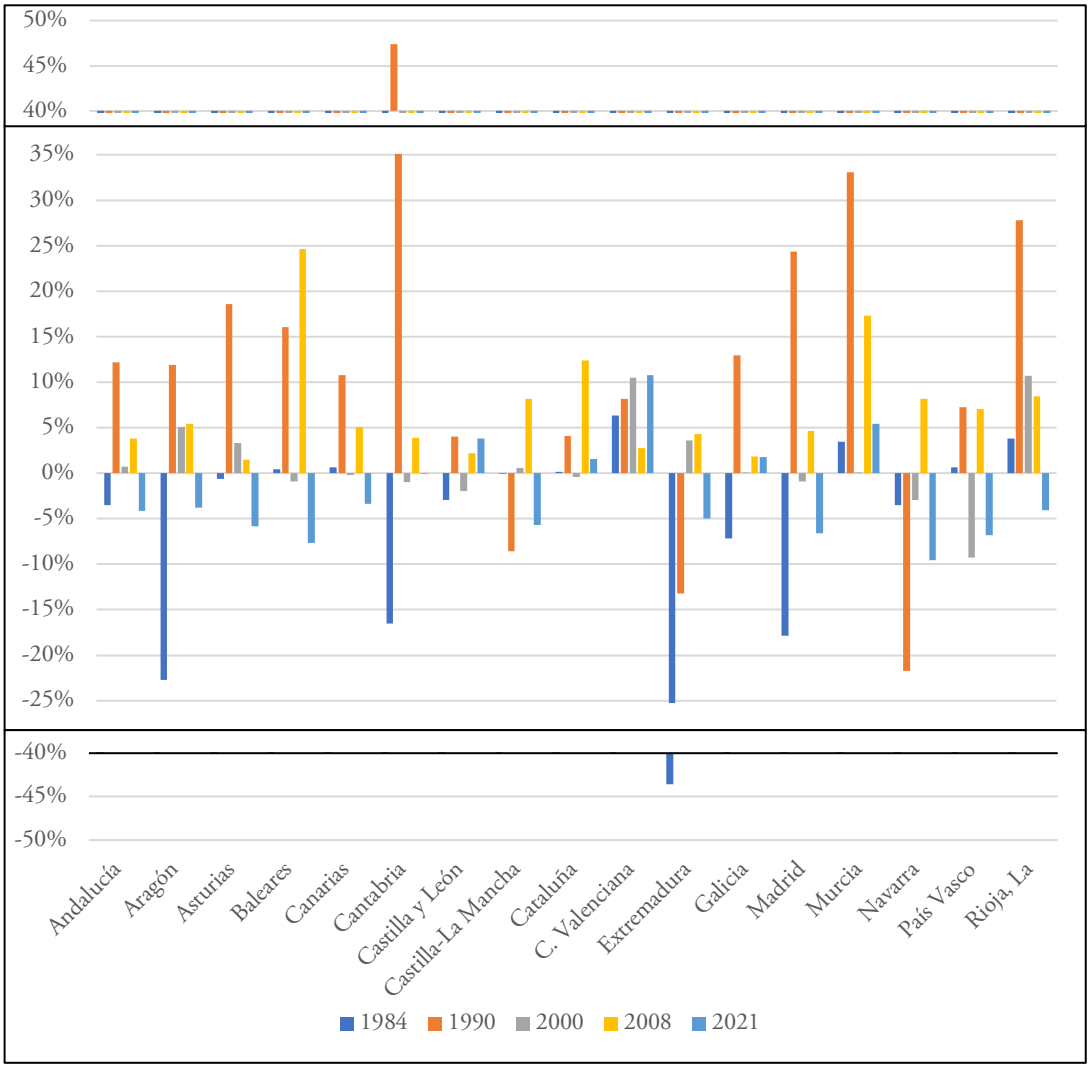


Fuente: Elaboración propia con datos del Ministerio de Hacienda.

La Figura 11 resume la evolución del gasto de capital y sus fuentes de financiación. Se observa un aumento significativo en el gasto de capital hasta 2008, seguido de un descenso sostenido hasta 2017, momento a partir del cual se inicia un incremento con tasas de crecimiento diversas entre regiones. La dinámica del ahorro revela dos patrones distintos: mientras que Navarra y Galicia mantienen un ahorro constante a lo largo del período, otras CC. AA enfrentan ciclos de ahorro y desahorro, con fluctuaciones más acentuadas en Castilla-La Mancha, Cataluña y Murcia. Las transferencias de capital también muestran una notable disminución a partir de 2007, especialmente en Asturias, Cantabria y Castilla y León. En contraste, algunas regiones, como Baleares y Cataluña, presentan variaciones menos pronunciadas. Desde 2017, se detecta un aumento en estas transferencias, particularmente en aquellas regiones que son beneficiarias de Fondos Europeos. La capacidad o necesidad de financiación osciló alrededor a cero hasta 2005, tras lo cual se incrementa, siendo más notable en Baleares, Cataluña y Murcia, mientras que Navarra y Galicia mantienen patrones más estables.

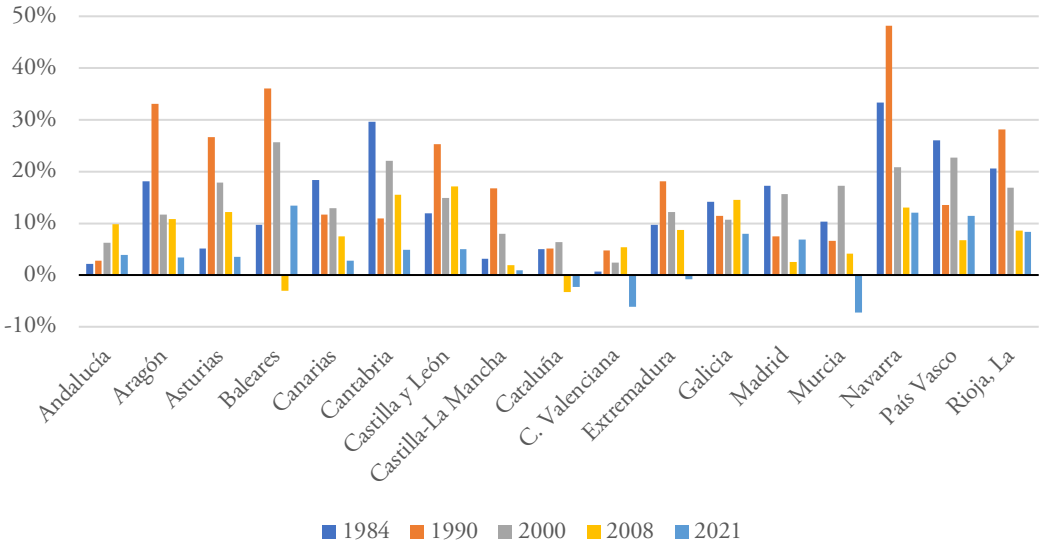
Las dinámicas observadas son el resultado de una variedad de factores, que incluyen las políticas de inversión pública, las condiciones económicas y las decisiones presupuestarias a nivel regional. Estas tendencias se reflejan en la figura 12, que ilustra la disminución del peso relativo del gasto en capital sobre el gasto no financiero regional desde la década de 1990.

FIGURA 9.
Déficit sobre ingresos no financieros en 1984,1990, 2000, 2008, y 2021. Datos en porcentaje



Nota: los valores negativos son superávit.
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

FIGURA 10.
Ahorro sobre ingresos no financieros,1984-1990-2000-2008-2021. Datos en porcentaje



Nota: los valores negativos son desahorro.
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

FIGURA 11.
Evolución del gasto de capital, las transferencias de capital, el déficit y el ahorro por CC.AA. (1984-2021)

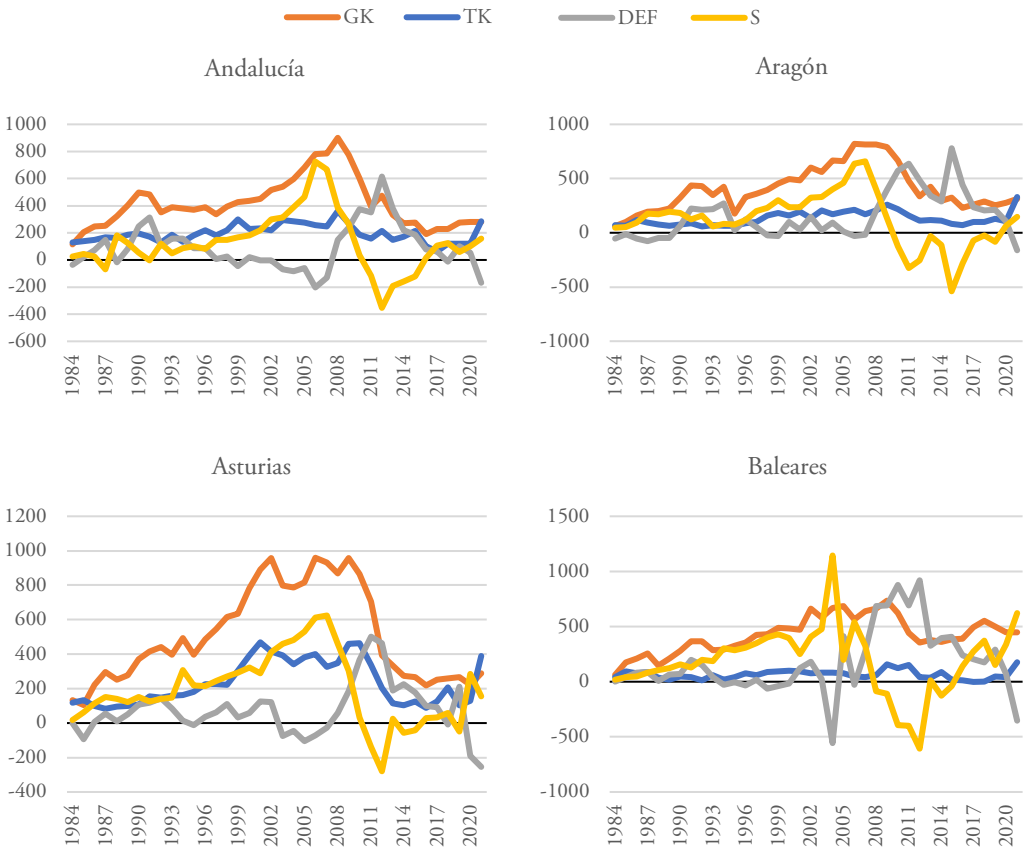


FIGURA 11. CONT.
Evolución del gasto de capital, las transferencias de capital, el déficit y el ahorro por CC.AA. (1984-2021)

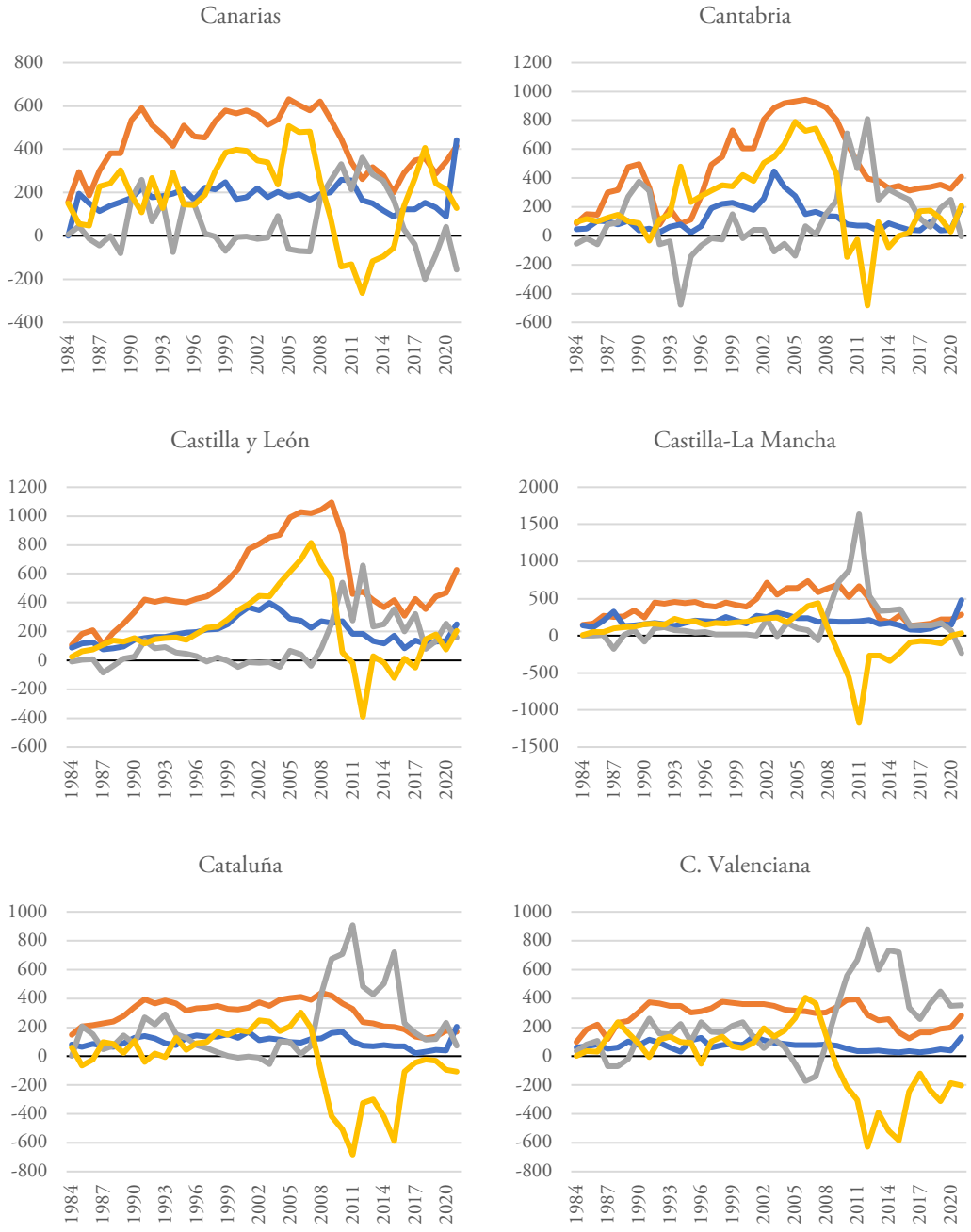


FIGURA 11. CONT.
Evolución del gasto de capital, las transferencias de capital, el déficit y el ahorro por CC.AA. (1984-2021)

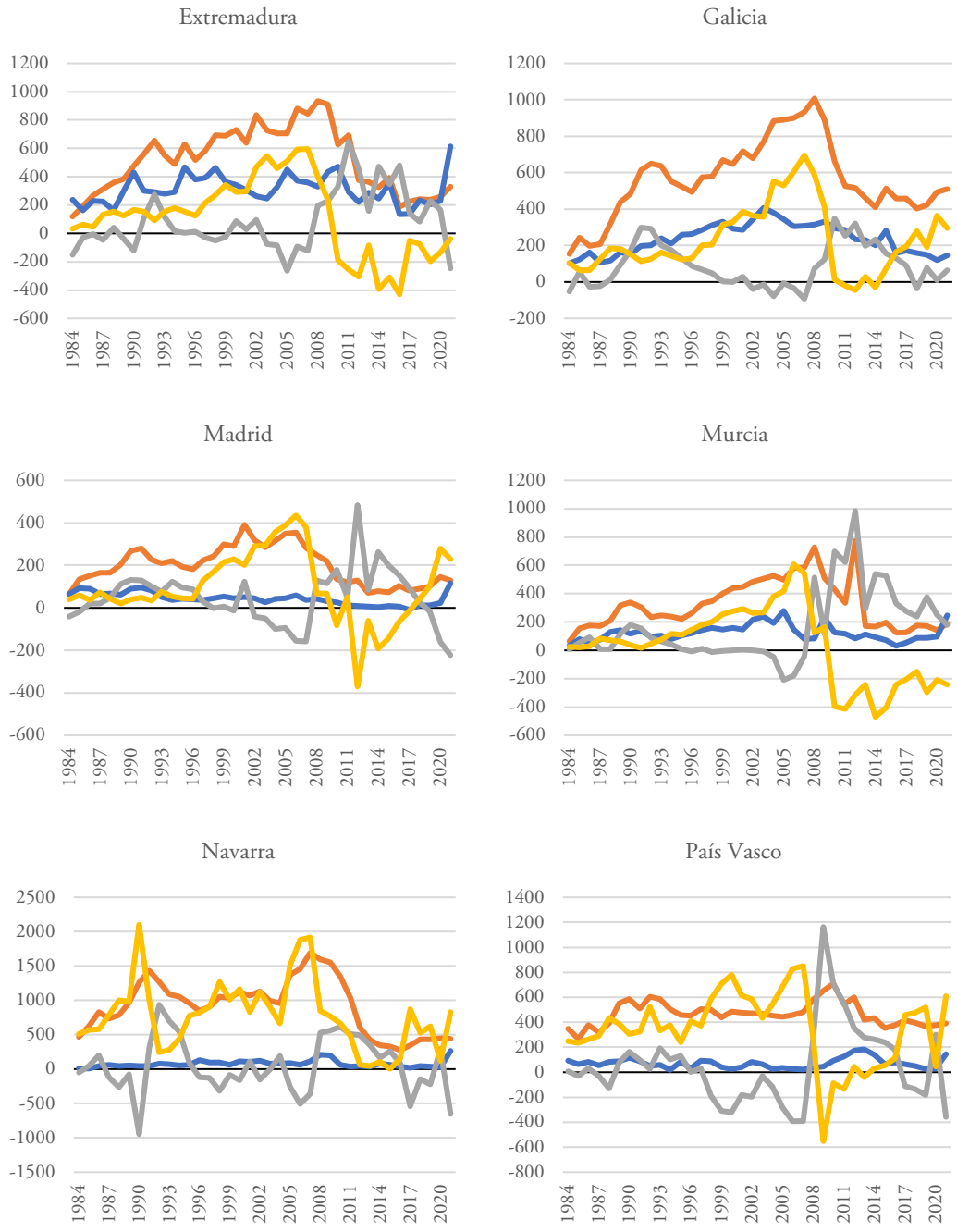
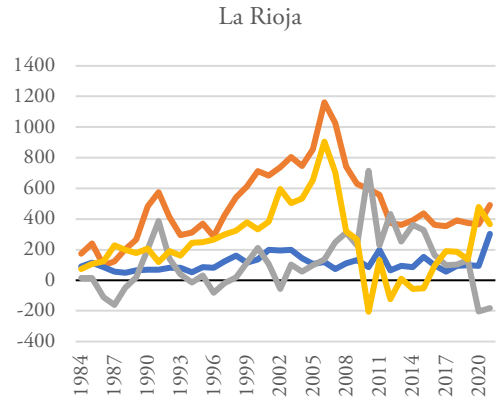
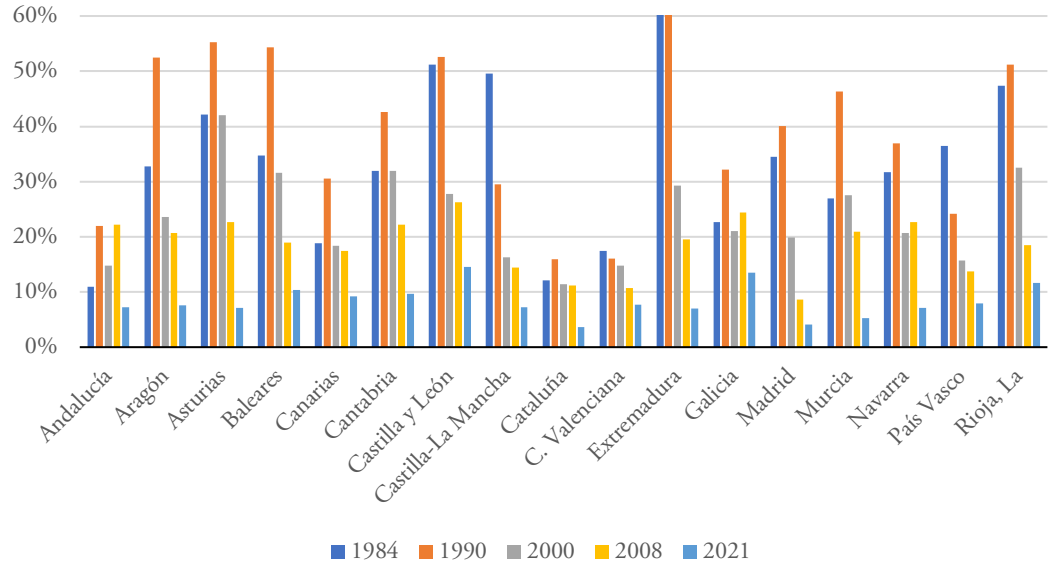


FIGURA 11. CONT.
Evolución del gasto de capital, las transferencias de capital, el déficit y el ahorro por CC.AA. (1984-2021)



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

FIGURA 12.
Gastos de capital sobre gastos no financieros, 1984-1990-2000-2008-2021. Datos en porcentaje



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

5. ANÁLISIS CLÚSTER Y σ-CONVERGENCIA

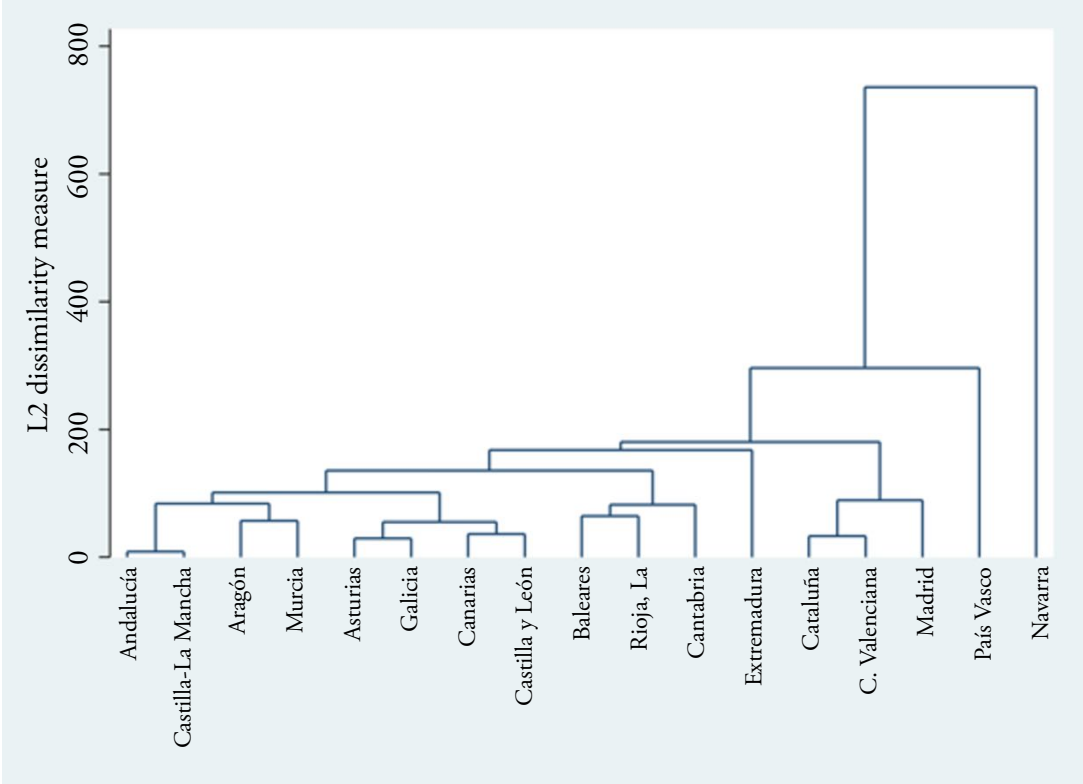
Para estructurar la compleja casuística de las CC. AA se ha realizado un análisis de clúster para examinar tanto la dinámica como los valores absolutos de las medias de las variables estudiadas. Los resultados se muestran en las Figuras 13 y 14. Complementariamente, se realiza un análisis σ-convergencia

para verificar si ha habido una reducción en la dispersión de los gastos y transferencias de capital per cápita entre CC.AA.[10]

La medida de disimilitud utilizada para el análisis de clúster es la distancia euclidiana, y se ha aplicado un algoritmo jerárquico basado en distancias medias. Posteriormente, se ha llevado a cabo una agrupación clúster utilizando el método de enlace-promedio (*average linkage*) para obtener una clasificación de las CC.AA.

En el primer subperíodo analizado, se identifican varios grupos claramente diferenciados. El más numeroso está compuesto por las quince comunidades de régimen común, que se dividen a su vez en cinco subgrupos: el primero incluye a Andalucía, Castilla-La Mancha, Aragón y Murcia; el segundo a Asturias, Galicia, Canarias y Castilla y León; el tercero a Baleares, La Rioja y Cantabria; el cuarto a Extremadura; y el quinto a Cataluña, Comunidad Valenciana y Comunidad de Madrid. Por otro lado, las Comunidades Forales de Navarra y el País Vasco forman un grupo independiente.

FIGURA 13.
Dendograma para las CC. AA en el período 1984-2007



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

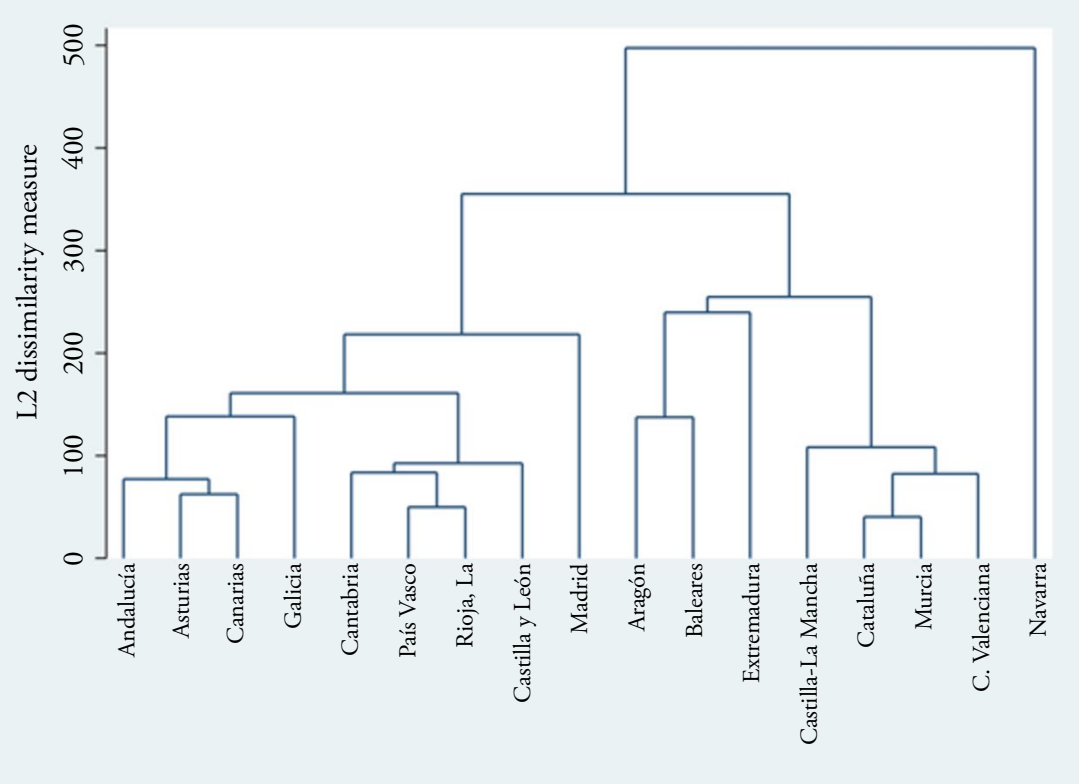
En el segundo subperíodo, Navarra sigue mostrando un comportamiento claramente diferenciado, mientras la distancia es menos pronunciada en las demás CC.AA. Para este análisis se han conformado dos grupos principales, de los que se excluye Navarra, un *outlier*. El primero se subdivide en tres subgrupos: Andalucía, Asturias, Canarias y Galicia; Cantabria, el País Vasco, La Rioja y Castilla y León; y la Comunidad de Madrid. El segundo conglomerado incluye tres subgrupos: Aragón y Baleares; Extremadura; y el tercero está compuesto por Castilla-La Mancha, Cataluña, Murcia y Comunidad Valenciana. A pesar de la existencia de varios grupos y subgrupos, las diferencias globales entre CC. AA se

[10] Se utiliza como indicador de σ-convergencia la desviación típica del logaritmo neperiano de los gastos (GK) y trasferencias de capital (TK) en términos per capita. No se considera el comportamiento del déficit (DEF) y ahorro (S) por los valores negativos registrados.

reducen en este período, en el que la medida de disimilitud resulta menor, lo que significa que los clústeres están más próximos entre sí.

Para completar esta evaluación, se realiza un análisis de σ-convergencia de la evolución de los gastos y transferencias de capital. Este enfoque permite evaluar hasta qué punto los valores de estas dos variables han convergido o divergido entre autonomías a lo largo del tiempo. Se considera que existe σ-convergencia cuando la desviación típica del logaritmo neperiano de las variables muestra una tendencia decreciente en el tiempo. Esto indicaría que las diferencias entre las CC. AA están disminuyendo, o lo que es lo mismo, se estaría registrando una homogeneización del gasto de capital y las transferencias del capital.

FIGURA 14.
Dendograma para las CC. AA en el período 2008-2021



Fuente: Elaboración propia con datos del Ministerio de Hacienda.

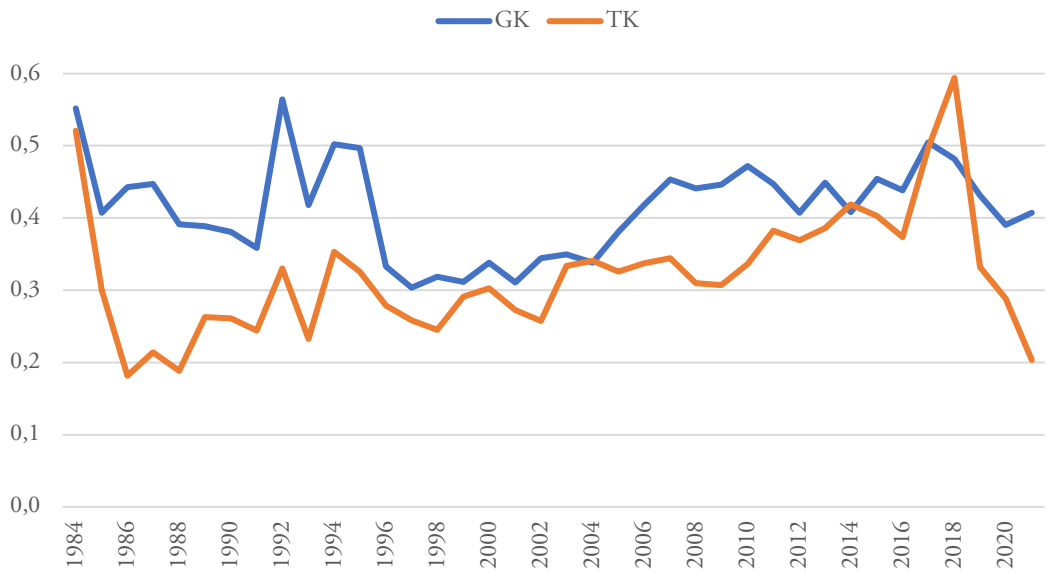
Como se observa en la Figura 15, durante el primer subperíodo analizado, la σ-convergencia para el gasto de capital pasó de 0,55 a 0,45, lo que indica un proceso de convergencia especialmente intenso entre 1996 y 2004, cuando los valores oscilaron entre 0,33 y 0,34. Sin embargo, en el segundo subperiodo, la dispersión volvió a aumentar, alcanzando el máximo de 0,50 en 2017, que representa el valor más alto de la serie.

Para las transferencias de capital, el proceso de σ-convergencia es aún más pronunciado que para el gasto de capital. Durante el primer subperiodo (1984-2007), el valor disminuye de 0,52 a 0,34, indicando un intenso proceso de convergencia durante el primer año. En el segundo subperíodo (2008-2021), el indicador descendió aún más, especialmente en los últimos años, cuando se produce una abrupta caída que puede ser explicada por las mayores dotaciones económicas para mitigar los efectos de la pandemia, especialmente desde la Administración General del Estado y la llegada de los Fondos Next Generation desde la Comisión Europea, que elevan sustancialmente y de manera similar los recursos extraordinarios transferidos a las CC.AA.

Del análisis efectuado, se derivan tres resultados principales. Primero, el comportamiento de la σ-convergencia para el gasto de capital y para las transferencias de capital está condicionado al ciclo

económico general. Normalmente, cuando se manifiesta una situación de recesión, se produce un recorte generalizado de los gastos de capital, lo que provoca que todas las CC. AA tengan importes similares en inversión. Lo anterior no es óbice para observar comportamientos dispares si algunas CC. AA, gracias a su situación económica saneada pueden permitirse la financiación del gasto de capital a través del ahorro-como sucede con las Comunidades Forales- o porque han recibido transferencias de capital para financiar este gasto. En segundo lugar, durante casi todo el período estudiado, el valor de la σ-convergencia de los gastos de capital ha sido superior al de las transferencias de capital. Tercero y último, la pandemia supuso un aumento notable de las transferencias en 2020 – 2021, incluyendo ya los *Next Generation* en todas las CC.AA. Esto explica la intensa convergencia tanto en las transferencias recibidas como en los gastos de capital.

FIGURA 15.
σ-convergencia de los gastos y las transferencias de capital (1984-2021)



Nota: Se calcula la σ-convergencia como la desviación típica del logaritmo neperiano de los gastos y transferencias per cápita.
Fuente: Elaboración propia con datos del Ministerio de Hacienda.

6. CONCLUSIONES

Las diferencias en los niveles de inversión entre las CC. AA están influenciadas por las fluctuaciones del ciclo económico y sus necesidades específicas. Durante las recesiones, es común que se priorice el gasto corriente en áreas esenciales como la sanidad, la educación y los servicios sociales, aunque esto plantea interrogantes sobre la sostenibilidad y eficiencia a largo plazo de esta estrategia.

A pesar de la recuperación económica, se observa una dificultad significativa para retornar a los esfuerzos inversores previos. En algunos casos, la inversión pública se mantiene significativamente por debajo de los niveles previos a la Gran Recesión, representando tan solo una tercera parte. Los patrones regionales varían según la dependencia de las CC. AA del FCI o de los Fondos Europeos, con comportamientos diferenciados en comunidades forales respecto a las del régimen común.

El análisis temporal revela un crecimiento notable en el gasto de capital hasta 2008, seguido de una disminución generalizada hasta 2017. La inversión pública estuvo condicionada por las reglas fiscales y su contribución al proceso de consolidación superó significativamente su peso relativo dentro del gasto total.

No obstante, la recuperación y el ajuste de la deuda y el déficit público trajeron consigo un repunte de la inversión, aunque con tasas de crecimiento variables entre regiones, según sus circunstancias.

El ahorro se erige como un factor crítico en la financiación de los gastos de capital. Navarra y el País Vasco destacan por su capacidad para generar ahorro, mientras que Cataluña y la Comunidad Valenciana representan el caso contrario. Por su parte, las transferencias de capital son un componente esencial para la financiación de las inversiones, especialmente en las CC. AA beneficiarias de Fondos Europeos, como Andalucía, Extremadura y Castilla-La Mancha.

Desde 2007, el incremento del déficit, que llegó a representar el 11,5% del PIB en 2009, ha generado una presión considerable sobre la capacidad de inversión de las CC.AA. Este desajuste entre ingresos y gastos ha conducido a una reducción de los recursos disponibles para financiar proyectos de infraestructura. Este fenómeno subraya la influencia del sistema de financiación regional y la capacidad de ahorro en el comportamiento del gasto en capital, con tendencias distintas en las CC. AA forales y las del régimen común.

Las conclusiones anteriores sugieren dos direcciones para futuras investigaciones. Primero, acometer un análisis más detallado del impacto específico de las subvenciones de capital en diferentes tipos de regiones dentro de la UE, en especial en las regiones Objetivo 1. Segundo, abordar la ampliación del horizonte temporal para poder incluir en los resultados la evaluación de las políticas recientes y el impacto de los fondos del programa *Next Generation* en las decisiones y el tipo de inversión, cuando se haya completado.

Agradecimientos

Agradecemos la ayuda de investigación de Alejandro Domínguez Lamela (GEN-Universidade de Vigo) y el apoyo financiero recibido por el Instituto de Estudios Fiscales a través del proyecto "Análisis de impacto de las transferencias de capital sobre la inversión y la deuda de las Comunidades Autónomas".

Referencias

Abiad, A., Furceri, D. y Tapalova, P. (2015). *The Macroeconomic Effects of Public Investment: Evidence from Advanced Economies.* (IMF Working Paper 15 (95)). 10.5089/9781475578874.001. https://www.imf.org/external/pubs/ft/wp/2015/wp1595.pdf

Alesina, A., Barbiero, O., Favero, C., Giavazzi, F. y Paradis, M. (2015). Austerity in 2009-2013. E*conomic Policy*, *30*, 383-437. https://doi.org/10.1093/epolic/eiv006

Alloza, M., Burriel. P. y Pérez. J.J. (2019). Fiscal policies in the euro área: Revisiting the size of spillovers, *Journal of Macroeconomics*. https://doi.org/10.1016/j.jmacro.2019.103132

Arrow, K. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, *29*(3), 155-173. https://doi.org/10.2307/2295952

Aschauder, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, *23*, 177-200. https://www.sciencedirect.com/science/article/abs/pii/0304393289900470?via%3Dihub

Asonuma, T. y Joo, H. (2020). Sovereign Debt Restructurings: Delay in Renegotiations and Risk Averse Creditors. *Journal of the European Economic Association, European Economic Association*, *18*(5), 2394-2440. http://dx.doi.org/10.2139/ssrn.2897145

Barro, R. (1991). Economic Growth in a Cross-Section of Countries. *The Quarterly Journal of Economics*, *106*(2), 407-443. https://doi.org/10.2307/2937943

Behrman, J. y Craig, S. (1987). The distribution of public services: an exploration of local government preferences. *The American Economic Review*, *77* (1), 37-49. https://www.sciencedirect.com/science/article/abs/pii/S0166046297000070

Bianchi, J., Ottonello, P. y Presno, I. (2019). *Fiscal Stimulus under Sovereign Risk*. (Working Papers 762, Federal Reserve Bank of Minneapolis). https://www.minneapolisfed.org/research/working-papers/fiscal-stimulus-under-sovereign-risk

Bosch, N. y Espasa, M. (1999). Con qué criterios invierte el sector público central? En Castells y Bosch (eds.) *Desequilibrios territoriales en España y en Europa*. Ariel.

Button, K. (1998). Infrastructure investment, endogenous growth and economic convergence. *The Annals of Regional Science*, *32*(1), 145-162. https://doi.org/10.1007/s001680050067

Case, A. (2001). Election goals and income redistribution: Recent evidence from Albania. *European Economic Review, 45*(3), 405-423. https://www.sciencedirect.com/science/article/abs/pii/S0014292100000787

Castells, A. y Solé-Ollé, A. (2005). The regional allocation of infrastructure investment: The role of equity, efficiency and political factors. *European Economic Review*, *49*(5), 1165-1205. https://www.sciencedirect.com/science/article/abs/pii/S0014292103001077

Comisión Europea (2024). *Eligibilidad para fondos estructurales.* https://ec.europa.eu/regional_policy/information-sources/maps_en

Conde-Ruíz J. I., Díaz M., Marín C. y Rubio J. (2016). Sanidad, Educación y Protección Social: Recortes Durante la Crisis. *Estudios sobre la Economía Española 2016/17*. FEDEA. https://documentos.fedea.net/pubs/eee/eee2016-17.pdf

De Jong, E., Ferdinandusse, J.M., Funda, J. y Vetlov, I. (2017). *The effect of public investment in Europe: a model based assessment.* (Documentos de Trabajo, 2021, Banco Central Europeo). https://www.ecb.europa.eu/pub/research/authors/profiles/josip-funda.es.html

De la Fuente, A. (2003). El impacto de los Fondos Estructurales: Convergencia real y cohesión interna. *Hacienda Pública Española*, *165*(2/2003), 122-129.

De la Fuente, A. (2013). Las finanzas autonómicas en boom y en crisis (2003-12). *Hacienda Pública Española/Review of Public Economics, 205*(2/2013), 127-150. https://www.ief.es/docs/destacados/publicaciones/revistas/hpe/205_Art5.pdf

Díaz Roldán, C. y Cuenca, A. (2014). *Consolidación fiscal y reglas fiscales en las Comunidades Autónomas españolas.* (Documento de trabajo, Catedra de Economía y Finanzas Internacionales). https://www.uv.es/~estevev/DT-E-2014-01.pdf

Fernández Leiceaga, X. y Lago Peñas, S. (2013). Sobre el reparto de los derechos del déficit entre las Comunidades Autónomas: una propuesta alternativa. *Revista de Economía Aplicada, 63*(XXI), 117-136. http://revecap.com/revista/numeros/63/pdf/fernandez_lago.pdf

Gordon, G. y Guerrón-Quintana, P. (2018). Dynamics of investment, debt, and default. *Review of Economic Dynamics*, *28*, 71-95. https://doi.org/10.1016/j.red.2017.07.007

Herrero-Alcalde, A., Martínez-Vázquez, J. y Murillo-García, E. (2011). Las transferencias de capital y la nivelación territorial. En S. Lago Peñas, y J. Martínez-Vázquez (Dirs.), *Las transferencias intergubernamentales en España. Análisis y propuestas de reforma*. Instituto de Estudios Fiscales.

Holtz-Eakin, D. (1994). Public-Sector Capital and the Productivity Puzzle. *The Review of Economics and Statistics*, *76*(1), 12-21. https://doi.org/10.2307/2109822

Holtz-Eakin, D. y Rosen, H. (1989). The "rationality" of municipal capital spending: Evidence from New Jersey. *Regional Science and Urban Economics*, *19*(3), 517-536. https://www.sciencedirect.com/science/article/abs/pii/0166046289900173

Instituto Nacional de Estadística (2024). Estadística continua de población (ECP). https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177095&menu=ultiDatos&idp=1254735572981

Lago-Peñas, S. (2015). Remaining changes to budgetary stability in Spain. *Spanish Economics and Financial Outlook (SEFO)*, *4*(2), 67-64. https://www.funcas.es/wp-content/uploads/Migracion/Articulos/FUNCAS_SEFO/018art08.pdf

Lago-Peñas, S., Fernández, X. Sánchez, P. y Álvarez, X. (2015). *Fondo de Compensación Interterritorial: análisis y propuestas de reforma.* Instituto Nacional de Administración Pública.

Lago-Peñas S., Fernández, X. y Vaquero, A. (2017a). ¿Por qué incumplen fiscalmente las Comunidades Autónomas? *Investigaciones Regionales - Journal of Regional Research*, *37*, 7-29. https://investigacionesregionales.org/es/article/por-que-incumplen-fiscalmente-las-comunidades-autonomas/

Lago-Peñas, S., Fernández-Leiceaga, X. y Vaquero, A. (2017b). Spanish Fiscal decentralization: A successful (but still unfinished) process. *Environment and planning C: and space*, *35*(8). https://doi.org/10.1177/2399654417704663

Lago-Peñas, S. y Vaquero, A. (2016). El comportamiento del déficit y la deuda en las Comunidades Autónomas en el período 2005-2015. *Mediterráneo Económico, 30*, 429-446.

Martínez-Vázquez, J., Muñoz, A. y Angarita, M. (2022). *El sistema de transferencias de capital en México: evaluación y propuestas de mejora.* (Documento IDB-DP-903. Banco Interamericano de Desarrollo). http://dx.doi.org/10.18235/0004194

Mas-Ivars, M., Pérez García, F. y Uriel, E. (2010): *El stock y los servicios del capital en España y su distribución territorial (1964-2008).* Fundación BBVA, Bilbao.

Mas-Ivars, M. y Pérez García, F. (2022). *El stock de capital en España y sus Comunidades Autónomas*. Fundación BBVA.

Mas-Ivars, M., Pérez, F., Uriel, E., Benages, E, y Cucarella, V. (2015). *Capital púbico en España. Evolución y distribución territorial (1900-2012), Informes 2015. Economía y Sociedad*. Fundación BBVA.

Mas-Ivars, M. y Pérez-García, F. (dirs) (2021). *El stock de capital en España y sus Comunidades Autónomas. Revisión metodológica y evolución reciente de la inversión y el capital 1995-2020*. Fundación BBVA.

Ministerio de Hacienda y Función Pública (2017). *Fondos de compensación interterritorial. Estadísticas e informes*. https://www.fondoseuropeos.hacienda.gob.es/sitios/dgfc/es-ES/ipr/oipr/fci/paginas/inicio.aspx

Ministerio de Hacienda (2024). *Liquidación de los Presupuestos de las Comunidades Autónomas.* https://serviciostelematicosext.hacienda.gob.es/SGCIEF/PublicacionLiquidaciones/aspx/menuInicio.aspx

Munell, A. H. (1992). Policy watch: infrastructure investment and economic growth. *The Journal of Economic Perspectives*, *6*(4), 189-198. https://www.aeaweb.org/articles?id=10.1257/jep.6.4.189

Murillo, E. y Sosvilla, S. (2005). Efectos de la oferta sobre la economía andaluza de las ayudas procedentes de los fondos estructurales destinadas a infraestructuras. El marco de apoyo comunitario 1994-1999. *Investigaciones Regionales*, *6*, 91-124. https://investigacionesregionales.org/es/article/efectos-de-oferta-sobre-la-economia-andaluza-de-las-ayudas-procedentes-de-los-fondos-estructurales-destinadas-a-infraestructuras/

Noveli, A. y Barcia, G. (2021). Sovereingn risk, public investiment and the fiscal policy stance. *Journal of Macroeconomics*, *67*. https://doi.org/10.1016/j.jmacro.2020.103263

Peña, A. (2008). Las disparidades económicas regionales en España: Las infraestructuras como factor de convergencia en el periodo 1980-2000. *Revista de Estudios Regionales*, *82*, 105-132. http://www.revistaestudiosregionales.com/documentos/articulos/pdf1042.pdf

Pastor, J. M., Pons, E. y Serrano, L. (2010). Regional inequality in Spain: Permanent income versus current income. *The Annals of Regional Science*, *44*(1), 121-145. https://doi.org/10.1007/s00168-008-0236-9

Pérez, J. J., y Solera. I. (2017). La evolución de la inversión pública durante la crisis y la recuperación. *Notas Económicas, Boletín Económico*, *4/2017*, Banco de España. https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/NotasEconomicas/T4/fich/bene1704-nec10.pdf

Ramey, V. A. (2020). The macroeconomic consequences of infraestructura investment. (NBER Working paper). https://www.nber.org/papers/w27625

Romer, P. (1986). Increasing returns and long run growth. *Journal of Political Economy*, *94*(5), 1002-1037. https:// www.jstor.org/stable/1833190

Ruiz-Huerta Carbonell, J. y García Díaz, M. A. (2012). El endeudamiento de las Comunidades Autónomas: Límites y problemas en el contexto de la crisis económica. *Revista d´estudis autonnómics i federals*, *15*, 124-163. https://raco.cat/index.php/REAF/article/view/252674

Solow, R. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, *70*(1), 65-94. https://doi.org/10.2307/1884513

ORCID

Alberto Vaquero García https://orcid.org/0000-0002-8697-4847

Santiago Lago Peñas https://orcid.org/0000-0003-4601-8655

María Cadaval Sampedro https://orcid.org/0000-0001-7948-6004

Patricio Sanchez Fernandez https://orcid.org/0000-0002-7848-3634

ANEXO

CUADRO 4.
CC. AA beneficiarias de los Fondos Europeos y del FCI (1984-2021)

	Fondos Europeos [1]			FCI		
	Desde	Hasta	Número de años	Desde	Hasta	Número de años
Andalucía	1989	2021	33	1984	2021	38
Aragón	0	0	0	1984	1989	6
Asturias	1989	2013	25	1984	2021	38
Baleares	0	0	0	1984	1989	6
Canarias	1989	2006	18	1984	2021	38
Cantabria	1994	2006	13	1984	2021	35
Castilla y León	1989	2006	18	1984	2021	38
Castilla-La Mancha	1989	2021	33	1984	2021	38
Cataluña	0	0	0	1984	1989	6
C. Valenciana	1989	2006	18	1984	2021	38
Extremadura	1989	2021	33	1984	2021	38
Galicia	1989	2013	25	1984	2021	38
Madrid	0	0	0	1984	1989	6
Murcia	1989	2021	33	1984	2021	38
Navarra	0	0	0	1984	1989	6

CUADRO 4. CONT.
CC. AA beneficiarias de los Fondos Europeos y del FCI (1984-2021)

	Fondos Europeos [1]			FCI		
	Desde	Hasta	Número de años	Desde	Hasta	Número de años
País Vasco	0	0	0	1984	1989	6
Rioja, La	0	0	0	1984	1989	6

Notas: (1) 1984-2006: regiones Objetivo 1; 2007-2013: regiones del objetivo convergencia, competitividad, *phashing in* y *phasing out;* 2008-2021: regiones menos desarrolladas, más desarrolladas y de transición.
Fuente: Elaboración propia con datos del Ministerio de Hacienda y Comisión Europea.

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 43-65
https://doi.org/10.38191/iirr-jorr.25.005

Articles

Convergence in the Spanish and Portuguese NUTS 3 regions: an analysis of the period 2000-2019

María Gema Flores Polán, Gemma Fuentes Calle**, Luís Fernando De la Macorra Cano****

Received: 12 April 2024
Accepted: 13 February 2025

Abstract:

The objective of this paper is to examine the convergence of economic activity between the various NUTS 3 regions of the Iberian Peninsula over the period 2000-2019. An analysis of the GDP (PPS) per capita of Spanish and Portuguese provinces was conducted with the help of different usual statistical and econometric methods of σ, β and γ. This was done to confirm the existence or non-existence of economic convergence in the period under analysis. The results revealed the existence of economic convergence between the Portuguese and Spanish provinces. In addition, they showed that more than half of the territory analysed was characterised by a scenario of poverty.

Keywords: Economic convergence; GDP; Spain; Portugal.
JEL Classification: O18; R11; R13.

Convergencia en las regiones NUTS 3 de España y Portugal: un análisis del periodo 2000-2019

Resumen:

El objetivo del presente trabajo es explorar la convergencia de la actividad económica entre las distintas regiones NUTS 3 de la Península Ibérica durante el periodo 2000-2019, a partir de un análisis del PIB (PPS) per cápita de provincias españolas y portuguesas, con ayuda de los diferentes métodos estadísticos y econométricos habituales de σ, β y γ. Para así poder confirmar la existencia o no de convergencia económica en el periodo analizado. Los resultados nos revelan la existencia de convergencia económica entre las provincias portuguesas y españolas. Además de mostrarnos un escenario de pobreza en más de la mitad del territorio analizado.

Palabras clave: Convergencia económica; PIB; España; Portugal.
Clasificación JEL: O18; R11; R13.

* Departamento de Economía Financiera y Contabilidad. Facultad de Ciencias Económicas y Empresariales. Universidad de Extremadura. España. gflores@unex.es
** Departamento de Economía Financiera y Contabilidad. Facultad de Administración y Dirección de Empresas. Universidad de Extremadura. España. gemmafu@unex.es
*** Departamento de Fundamentos del Análisis Económico. Facultad de Ciencias Económicas y Empresariales. Universidad de Extremadura. España. lfmacorra@gmail.com
Corresponding author: gflores@unex.es

1. Introduction

There is significant discussion about inequalities between countries, between rich and poor countries. These inequalities are said to be the result of globalisation and the free market economy. However, some studies show that this is a mirage. Analyses of lower-income countries' economies show that differences between countries and income inequalities are decreasing over time. The process of economic growth has affected the countries with the poorest citizens in the world, the differences between rich and poor have significantly decreased in recent years (Pinkovskiy *et al*., 2024; Sala-i-Martin, 2012). This is known as economic convergence, when a poor economy grows faster than a rich economy and can catch up in terms of per capita income (Barro and Sala-i-Martin, 2009).

Nevertheless, the theories of rich and poor individuals are still valid today. The economic gap exists if we focus on analysing the difference in per capita income of individuals in the same country, that is, the difference in income between citizens indicates the presence of both rich and less rich individuals within the same territory (Sala-i-Martin, 2012). Based on the above, we can discuss the coexistence of prosperous and depressed areas within the same geographical regions. According to these ideas, it has been verified that there are differences in economic growth between various regions in Europe (Villaverde and Pérez, 1996).

It is important to acknowledge that the elimination of cross-border barriers within the European Union (EU) has exposed substantial economic, social, and territorial disparities among member states. It was therefore, expected that the opening of borders would lead to convergence, thereby helping to reduce these economic inequalities. Consequently, some authors claim that cohesion between regions is achieved through economic convergence, which is one of the main objectives of EU policies (López-Villuendas and del Campo, 2022; Pires and Nunes, 2018). The EU promotes relations and cooperation between cross-border territories, which is why these areas are considered "laboratories of European integration" (European Commission, 2021; López-Villuendas and del Campo, 2022).

In this context, the aim of this paper is to examine economic convergence in the Iberian Peninsula and in the so-called Iberian south-west corridor over the period 2000-2019. This objective is achieved through an analysis of GDP (PPS) per capita in Spanish and Portuguese NUTS 3 regions, using different methods to calculate the convergence of economic activity.

As a justification for this research, it is important to note that the study of growth theories and problems related to economic growth and development is currently reflected in numerous research papers. These studies attempt to explain the causes of economic growth between countries or regions through the analysis of convergence (León, 2013). However, our analysis focuses on the NUTS 3 regions of Spain and Portugal, as these countries share a border and joined the EU in the same year. Despite their structural differences, they have managed to achieve the objectives of the European and Monetary Union (Viegas and Antunes, 2013).

Furthermore, in the case of Spain and Portugal, there are not many empirical analyses that analyse the existence of convergence between these two neighbouring countries. Moreover, the analyses that have been carried out refer to earlier periods, between 1955 and 2008. The results of economic convergence between Spain and Portugal differ depending on the period analysed, despite the economic growth of both areas, the existence of economic convergence in these territories should be the aim to be achieved in the future (López Martínez, 2001; López-Villuendas and del Campo, 2022).

An interesting and recent study for Spain and Portugal is that of Villuendas and Del Campo (2022). They analyse the convergence between the Galicia-Northern Portugal Euroregion over the period 1980-2019 and determine the absolute convergence between the two territories. However, no study has been carried out for the south-western territories of the Iberian Peninsula, which is why we consider it interesting to do so in this work. In addition, we carry out the analysis of the Iberian southwest with and without capitals, in order to eliminate the synergy of the most developed regions.

Through our empirical analysis, it has been observed that, over the period 2000-2019, the economies of the NUTS 3 regions of the Iberian Peninsula have undergone a gradual process of improvement in the

degree of sigma convergence, in terms of GDP (PPS) per capita. However, as far as γ convergence is concerned, a different behaviour can be observed, with neither absolute stability nor complete destabilisation of the ranking variation. Thus, important internal asymmetries persist and need to be addressed in order to achieve more equitable and sustainable convergence.

This research work serves to confirm the existence of economic convergence in the Iberian Peninsula and the Iberian Southwest Corridor during a period characterised by the advent of the 21st century and the concomitant experience of diverse economic scenarios encompassing crises and recoveries. Considering the evidence presented, it can be concluded that the hypothesis of economic convergence between the Spanish and Portuguese provinces is accepted. However, it is notable that there is a significant disparity in per capita income levels across the Peninsula, with persistent differences in income levels between regions.

The rest of the article is structured by headings as follows. Following this introduction, point (2) presents the theoretical framework. Point (3) describes the methodology used based on the available data. In point (4), the results are presented and discussed. Finally, in point (5), we present the conclusions.

2. Theoretical framework

Since the 1930s, the effects of economic growth have been studied, leading to the emergence of the term "economic convergence". Researchers have been interested in explaining the factors that influence economic development in various territories (León, 2013). Studies by Solow and Swan (1956), Baumol (1986), Baumol and Wolf (1988), and Barro and Sala-i-Martín (1990, 1991, 1992) are considered the most influential in the analysis of economic convergence at the international level.

In the last decade of the twentieth century, Sala-i-Martín (1990) was the first to use the convergence techniques that have become the most widely used by regional analysts: beta convergence (β) and sigma convergence (σ) (Villaverde, 2004).

σ-Convergence analyses the evolution over time of the dispersion observed in the reference variable. The two basic measures of dispersion are the standard deviation of the logarithms and the coefficient of variation.

β-Convergence compares the change between the starting point of the period and the current situation. This process is called unconditional or absolute convergence, and beta convergence occurs when this relationship is negative. In terms of economic growth, the beta convergence hypothesis suggests that regions that start with a lower per capita income experience higher growth rates. This technique involves estimating the relationship between the growth of a variable in each period and the value of that variable at the starting point.

It should be noted that while σ convergence is a sufficient condition for beta convergence, it is not a necessary one. Boyle and McCarthy (1997, 1999) proposed a new concept of convergence to be used together with sigma to test for the existence of beta, called gamma convergence (ɣ). The rationale behind this proposal is based on the premise that approximation or convergence is a more comprehensive concept than the previous definitions. It is therefore evident that the measures must be based on the reduction of dispersion (σ) or the increase of smaller observations (β) and must also account for movements in the distribution that alter the ranking over time (γ). This factor has not been considered in the two measures previously presented.

In the specific case of Spain and Portugal, we find some studies analyzing the convergence of these countries. Rodríguez (1999), through the analysis of GDP per capita in the Spanish regions during the period 1955-1996, shows that the causes of regional convergence were, on the one hand, the movements produced in the productive factors. Specifically, they highlight the population and, on the other hand, the convergence of productive sectoral structures, which contributed to the growth of average regional productivity equality. Nevertheless, this positive evolution of regional equality has slowed down since 1979 due to population migration caused by increasing unemployment rates in different regions. Additionally, the disparities between regions are due to differences in productivity and employment rates. For this reason,

we proceed to the factorial decomposition of GDP (PPS) per capita into its two components (productivity and an approximation of the employment rate) for further analysis.

On the other hand, Peres (2004) analyses labour productivity at the sectoral level in the NUTS 3 regions of continental Portugal during the 1990s. Again, the author finds that labour mobility favours regional convergence in productivity. However, capital accumulation does not contribute to regional convergence.

In addition, Villaverde (2004) states that it is necessary to carry out a study of economic convergence at the provincial level, since the existing studies in Spain take the Autonomous Community as the unit of analysis. This author tries to overcome the problem of what Sala-i-Martín (1996) calls the "classical approach to convergence", which considers spatial units as if they were independent of each other. This work shows the convergence observed in Spain during the period 1985-2002. By analysing the variable of the productivity of the labour factor, he concludes that there is convergence β and σ, observing that there is a concentration of GDP and employment in the most efficient provinces. There is more intensity in the employment variable than in GDP.

Other studies show that during the period 1985-1999 there was no economic convergence between the different EU countries, due to the different situations in each country, so there was no reduction in inequality in the EU during this period. Focusing on Spain, at the beginning of the period (1985) the ranking of the Autonomous Communities was led by Madrid and the Balearic Islands with a higher GDP pc. This was not the case for the Autonomous Communities of Extremadura and Andalusia, which were at the bottom of the list with the lowest GDP pc in Spain (below the average). This situation changed by the end of the period (1999), when Catalonia joined Madrid at the top of the list. By this time, the autonomous communities of Extremadura, the two Castilian communities and Galicia had improved their situation and were above the average.

Viegas and Antunes (2013) contribute to the study of the economic convergence in the Spanish and Portuguese NUTS 3. The authors conclude that these two areas have experienced economic growth during the period 1985-2008 and converged towards European per capita income levels. However, the analysis does not show economic convergence between the poorer regions and concludes the existence of economic clusters.

A further study that analyses β and σ convergence, this time for the regions of Galicia (Spain) and Northern Portugal, is by López-Villuendas and Del Campo (2022). They analyse whether cooperation between the two cross-border regions contributes to economic convergence in both areas during the period 1980-2019. Their results indicate the existence of economic convergence throughout the period, in addition to the existence of absolute convergence. Thus, it is shown that cooperation between these regions has led to a process of cohesion.

Turning now to the south-west Iberian corridor, it is worth noting the scarcity of studies on this subject. According to López-Villuendas and Del Campo (2022), the results of their analysis of economic convergence between Galicia and northern Portugal can be extrapolated to other EU Euroregions to explain existing disparities and understand the process of convergence between cross-border regions.

In this line, we define the Iberian Southwest Corridor as the new axis of European development (García and Mora, 2022). Through their work, García and Mora (2022) aim to highlight an instrument that, in their opinion, will be able to define a new scenario in the European and Iberian context. This south-western Iberian corridor does not currently exist because there is no infrastructure to connect it. It is evident that the project has the potential to serve as a conduit for articulating the Iberian Peninsula, connecting the two capital cities, Madrid and Lisbon, and the regions between them. This is what these authors call "terra ignota", which today, in the 21st century, can be understood as ignored land. Ignored land, without infrastructure, condemned to a vicious cycle of poverty. These are territories that belong to a prosperous Europe, yet do not achieve the development towards progress and well-being (López-Villuendas and del Campo, 2022).

The Iberian southwest can be understood as the area that coincides with the basins of the Tajo and Guadiana rivers as they flow towards the Atlantic. This area, which connects the centre of the Iberian

Peninsula with the Atlantic coast, was known as Lusitania before the Roman era. In the 13th century, these territories were divided, and the oldest border in Europe appeared, dividing the Iberian Peninsula into two territories that had been at odds for years. The establishment of the border and the lack of communication infrastructure contributed to the creation of two isolated territories, which are undoubtedly the origin of the economic and social conditions in the regions of Alentejo in Portugal and Extremadura in Spain. This isolation and lack of communication have been the cause of the culture of survival observed in the Raya (the border between the two territories) over the centuries (García, 2022).

3. Methods

3.1. Data description

The indicator used to observe the economic differences between the different provinces and to analyse economic convergence is the GDP (PPS) per capita, taken from the database of the National Institute of Statistics of Spain and Portugal for the period 2000-2019, for which we found the most recent data available. The choice of this interval is due to its lower limit being the beginning of the century and its upper limit being the year before the Covid pandemic. The sample consists of 84 provinces of the Iberian Peninsula, including the islands and the autonomous cities of Ceuta and Melilla.

First, a descriptive analysis of GDP per capita per capita is conducted using the average of the Iberian Peninsula and the Iberian Southwest. Secondly, based on the factorial decomposition of GDP (PPS) p.c. into its two components, GDP per person employed (productivity) and employment in the population (an approximation of the employment rate), the analysis of the cumulative annual average rates of change (hereafter CAAR) is conducted for the variables GDP per capita, employment, and productivity of the Iberian Peninsula and the Iberian Southwest, with and without capitals (Madrid and Lisbon). To do this, we analyse the changes that have taken place over the period analysed, bearing in mind that more developed regions, such as Madrid or Lisbon, may have different development patterns compared to less developed areas.

Regarding the period analysed, 2000-2019, it has been divided into three sub-periods based on recent historical and economic events in the world and in the Iberian Peninsula, in order to observe the evolution of GDP (PPS) per capita. For this purpose, we take as a reference the beginning of the century (2000-2007), the great international economic and financial depression of 2008 (2008-2014) and the beginning of the recovery until one year before the generalisation of the Covid pandemic (2015-2019).

To analyse the data, we have selected the NUTS 3 regions that make up the Iberian Peninsula (Spain and Portugal) and the NUTS 3 regions that make up the Southwest Iberian Corridor (Tables 1 y 2). As a justification for the use of NUTS 3, the European Union has made the NUTS 3 nomenclature available to researchers to facilitate the collection and transmission of harmonized regional statistics, due to the need for comparable data. The use of NUTS 3 analysis is justified by the need to ensure the comparability of the regions of Spain and Portugal. The NUTS 3 nomenclature makes it possible to compare the geographical units in terms of population, political, administrative and institutional reality. In addition, this nomenclature reflects economic, social, historical, cultural, geographical, and environmental factors (Gouardères 2024).

On the order hand, as justification for the sample of the Southwest Iberian Corridor, it must be noted that the EU single market requires the elimination of all physical and geographical barriers to facilitate the exchange of people and goods throughout the European territory (Fernandez, 2022). The "terra ignota", as the territories of the Iberian southwest are known, are territories belonging to developed Europe. Yet, they are territories that do not achieve the development and welfare of the rest of Europe today; in addition, they are considered territories condemned to poverty (López-Villuendas and del Campo, 2022). Therefore, it is interesting to observe whether these territories, isolated by the border and the lack of communication infrastructures, have evolved throughout the period analyzed due to the effect of economic convergence. In this way, it will be possible to attract the attention of the authorities and

secure the necessary support to promote the recovery of the two nations (López-Villuendas and del Campo, 2022).

Table 1.
Variables, sources and descriptives

Variable	GDP PPS
Definition	Gross Domestic Product (GDP) is a measure of the total output of goods and services produced by an economy. It is calculated by subtracting intermediate consumption and adding net taxes on products and imports. The GDP per capita is derived by dividing the GDP by the average population for a given year. Purchasing Power Parities (PPPs) serve as a common currency that neutralises price level differences across countries, facilitating meaningful comparisons of GDP volumes.
Sources	National Institute of Statistics of Spain and Portugal
Iberian NUTS 3 Average	20.445,96€
Iberian NUTS 3 Standard desviation	4.766,61€
Southwest area Average	19.924,17€
Southwest area Standard desviation	5.155,76€
Southwest area Average without capital	17.951,00€
Southwest area Standard desviation without capital	3.541,58€
Variable	**EMPLOYMENT (thousands of people)**
Definition	The employed population is defined as all persons of working age who are engaged in paid employment.
Sources	National Institute of Statistics of Spain and Portugal
Iberian NUTS 3 Average	287,75
Iberian NUTS 3 Standard desviation	461,18
Southwest area Average	477,61
Southwest area Standard desviation	882,01
Southwest area Average without capital	121,04
Southwest area Standard desviation without capital	71,24

TABLE 1. CONT.
Variables, sources and descriptives

Variable	POPULATION (thousands of people)
Definition	Population is defined as the set of people who habitually reside in a given territory.
Sources	National Institute of Statistics of Spain and Portugal
Iberian NUTS 3 Average	659,83
Iberian NUTS 3 Standard desviation	953,54
Southwest area Average	1.010,52
Southwest area Standard desviation	**1.707,86**
Southwest area Average without capital	319,30
Southwest area Standard desviation without capital	213,94

Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

TABLE 2.
NUTS 3 regions of the Iberian Peninsula (NUTS 2024 classification)

Code	Region	Code	Region
ES111	A Coruña	**ES612**	Cádiz
ES112	Lugo	**ES613**	Córdoba
ES113	Ourense	**ES614**	Granada
ES114	Pontevedra	**ES615**	Huelva
ES120	Asturias	**ES616**	Jaén
ES130	Cantabria	**ES617**	Málaga
ES211	Álava	**ES618**	Sevilla
ES212	Gipuzkoa	**ES620**	Murcia
ES213	Vizcaya	**ES630**	Ceuta
ES220	Navarra	**ES640**	Melilla
ES230	La Rioja	**ES703**	El Hierro
ES241	Huesca	**ES704**	Fuerteventura
ES242	Teruel	**ES705**	Gran Canaria
ES243	Zaragoza	**ES706**	La Gomera
ES300	Madrid	**ES707**	La Palma
ES411	Ávila	**ES708**	Lanzarote
ES412	Burgos	**ES709**	Tenerife
ES413	León	**PT111**	Alto Minho
ES414	Palencia	**PT112**	Cávado
ES415	Salamanca	**PT119**	Ave

TABLE 2. CONT.
NUTS 3 regions of the Iberian Peninsula (NUTS 2024 classification)

Code	Region	Code	Region
ES416	Segovia	**PT11A**	Á.M.do Porto
ES417	Soria	**PT11B**	Alto Tâmega
ES418	Valladolid	**PT11C**	Tâmega e Sousa
ES419	Zamora	**PT11D**	Douro
ES421	Albacete	**PT11E**	T. de T. Montes
ES422	Ciudad Real	**PT150**	Algarve
ES423	Cuenca	**PT1D1**	Oeste
ES424	Guadalajara	**PT191**	Região de Aveiro
ES425	Toledo	**PT192**	Região de Coimbra
ES431	Badajoz	**PT193**	Região de Leiria
ES432	Cáceres	**PT194**	Viseu Dão Lafões
ES511	Barcelona	**PT195**	Beira Baixa
ES512	Girona	**PT1D2**	Médio Tejo
ES513	Lleida	**PT196**	B. e S. da Estrela
ES514	Tarragona	**PT1A0**	Á. M. de Lisboa
ES521	Alicante	**PT1C1**	Alentejo Litoral
ES522	Castellón	**PT1C2**	Baixo Alentejo
ES523	Valencia	**PT1D3**	Lezíria do Tejo
ES531	Ibiza	**PT1C3**	Alto Alentejo
ES532	Mallorca	**PT1C4**	Alentejo Central
ES533	Menorca	**PT200**	R. A. dos Açores
ES611	Almería	**PT300**	R. A. da Madeira

Source: Own elaboration using data from Eurostat.

TABLE 3.
NUTS 3 Regions of the Iberian Southwest

Code	Region
ES300	Madrid
ES425	Toledo
ES432	Cáceres
ES431	Badajoz
ES422	Ciudad Real
PT195	Beira Baixa
PT1D2	Medio Tajo
PT1D3	Leziria do Tajo
PT1C3	Alto Alentejo
PT1C4	Alentejo Central
PT1C1	Alentejo Litoral
PT1A0	A.M. Lisboa

Source: Own elaboration using data from Eurostat.

3.2. Convergence

The objective of this study is to examine the economic convergence between the various provincial NUTS 3 spatial demarcations in Spain and Portugal. A second convergence analysis will be conducted using a sample of the provinces that define the Iberian Southwest corridor (García Salas and Mora Aliseda, 2022).

This analysis will employ a spatial demarcation based on Spanish and Portuguese NUTS 3 provincial units. The convergence of economic activity will be calculated using various common statistical and econometric techniques, including those based on σ, β, and γ.

The use of σ convergence entails the examination of the standard deviation of the logarithm of per capita income within a group of countries or regions. This is done to ascertain whether the dispersion diminishes over time, thereby indicating convergence. The standard deviation of the logarithms of GPD per capita has been calculated using Equation 1:

Equation 1.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}}$$

Sigma convergence, denoted by 'σ', is calculated as the standard deviation of GDP per capita. The logarithms of GDP per capita have been considered for each of the NUTS 3 regions in each year of the period analysed. Moreover, $\bar{x}$ represents the average of the logarithms of GDP per capita for the entire period from 2000 to 2019. Finally, N denotes the total number of NUTS 3 regions analysed.

Conversely, if the objective is to determine whether a less affluent economy tends to exhibit greater growth than a more prosperous one, β convergence is employed. This phenomenon is evidenced by a negative correlation between GDP per capita growth and the initial per capita income level. In other words, beta-convergence entails estimating the relationship between the growth of a variable over a given period and the value of that variable at the outset of the period in question (Sala-i-Martín, 1990).

The calculation of Beta convergence is achieved by means of the equation 2:

Equation 2.

$$y = log2019 - log2000$$

Beta convergence, denoted by 'y', is calculated as the difference between the logarithm of GPD per capita in the year 2019 and the logarithm of GPD per capita in the year 2000.

The hierarchical relationship between these two concepts, σ and β, is expressed as follows:

- The phenomenon of β convergence is defined as the tendency of economies with lower initial GDP per capita to exhibit higher rates of growth in comparison to those with higher initial GDP per capita.
- Conversely, σ convergence signifies the diminution in the dispersion in GDP per capita levels between economies over time.

Ultimately, to ascertain the presence of beta-convergence, a gamma-convergence analysis is conducted to determine whether shifts in the distribution within the specified period have influenced the ranking over time. The convergence γ has been calculated using Spearman's Rho for the GDP per capita variable. Equation 3 quantifies the Spearman rank correlation coefficient, also referred to as Spearman's Rho (ρ). This coefficient is used to analyse the monotonic relationship between two ordinal or continuous variables.

EQUATION 3.

$$r = 1 - (6 * \sum D^2 / n(n^2 - 1)$$

The Spearman correlation coefficient, denoted by r, is used to quantify the strength of the linear relationship between two variables. The term 'D' represents the difference between the ranks assigned to each pair of observations of the two variables. The term 'n' denotes the total number of data pairs in the sample, while '$\sum D^2$' signifies the sum of the squares of the differences in ranks.

The calculation of this coefficient is performed in the following manner:

1. First, the values of each variable are ordered separately.
2. Second, the values are assigned ranks, with 1 assigned to the lowest value, 2 assigned to the next lowest value, and so on.
3. Third, the differences (D) between the ranks for each pair of observations are calculated. These differences are squared and then summed.
4. Finally, the formula indicated in Equation 3 is applied.

The value of r will range between -1 and +1, where r = +1 indicates a perfect positive correlation, r = -1 indicates a perfect negative correlation, and r = 0 indicates no monotonic correlation between the variables

4. RESULTS AND DISCUSSION

4.1. GDP (PPS) PC

The following section presents a descriptive analysis of GDP (PPS) per capita, with the sample classified according to economic level. To this end, the sample of provinces has been divided into four groups based on their level of GDP per capita. Subsequently, the mean values of GDP per capita have been calculated for the period 2000-2019 and for each of the sub-periods into which the sample has been divided. Thus, it can be observed that, in general, GDP per capita has increased between the sub-periods considered within this extensive time frame for both the Iberian Peninsula and the Iberian Southwest (see tables 3 and 4). The mean value of GDP (PPS) per capita for the period 2000-2019 is €22,883.19 per inhabitant.

As previously stated, the provinces have been classified into four levels, delineated by the lowest and highest average values observed in the sample. Four intervals have been formed, comprising two above and two below the average value. This allows us to distinguish between provinces with low-incomes (€11,570-€17,227), provinces with lower middle-incomes (€17,227-€22,883), and provinces with upper-middle-incomes (€22,883- The final category comprises provinces with high-incomes (€28,816-€34,750), according to a classification system used by the World Bank to categorise countries according to income levels (Hamadeh et al., 2023).

In the sample analysed, there are eighteen low-income provinces, forty-four lower-middle-income provinces, eighteen upper-middle-income provinces and only four high-income provinces. This study will present a discussion of the most significant findings, with particular attention paid to the intervals whose values exceed the average (the third and fourth levels).

The analysis of the sample revealed the existence of eighteen provinces with low-incomes, forty-four provinces with low average incomes, eighteen provinces with high average incomes and, finally, only four provinces with high-incomes. The results of the provincies classified in each interval are presented in the following section.

The majority of provinces with low-incomes are located in Portugal, with the exceptions of Algarve, Leiria, Aveiro, Madeira, and the Porto Metropolitan Area, which are situated within the second interval

due to their low middle incomes. Furthermore, the first interval comprises the provinces of Cáceres, Badajoz, Granada, Córdoba, and Jaén. This classification is in accordance with the findings of Viegas and Antunez (2013) in their analysis of the NUTS 3 regions of Spain and Portugal for the period 1995-2008. The authors in question initiated their analysis with a minimum income of €7,435.12 and a maximum of €27,624.59. It is evident that the sample employed in our analysis commences with a minimum income of €11,570, which would be above the first interval utilised by the authors.

In the second interval, in addition to the Portuguese provinces already mentioned, the Spanish provinces with the lowest average incomes are: A Coruña, Lugo, Ourense, Pontevedra, Ávila, Burgos, León, Salamanca, Segovia, Zamora, Albacete, Ciudad Real, Cuenca, Guadalajara, Toledo, Alicante, Valencia, Cantabria, Asturias, Murcia, the autonomous cities of Ceuta and Melilla, Mallorca, Menorca, Huelva, Seville, Cádiz, Málaga, Almería, El Hierro, Fuerteventura, Gran Canaria, La Gomera, La Palma, and Tenerife. Viegas and Antunes (2013) proposed that their third and fourth levels should encompass the aforementioned provinces (excluding the islands and the autonomous cities of Ceuta and Melilla), with the proviso that differing income levels should be taken into account. The authors situate these provinces within the third level, delineating upper and lower limits of €13,938.61 and €17,836.91, respectively. However, our analysis indicates that income levels in these NUTS 3 regions are higher.

In the third interval, the provinces are ranked according to two distinct criteria. The first criterion is the ranking of the provinces with the highest GDP per capita on an individual basis. The second criterion is the ranking of the provinces with a high GDP per capita due to their low population density. The initial classification comprises Vizcaya, Barcelona, La Rioja, and Lisbon. In accordance with the aforementioned criteria, the following provinces are characterised by low population density: Huesca, Teruel, Zaragoza, Burgos, Palencia, Soria, Valladolid, Girona, Lleida, Tarragona, Castellón, Ibiza and Formentera and Lanzarote. Additionally, the Portuguese provinces of Lisbon and the Alentejo Litoral are included in this classification. It should be noted that the minimum income considered in this interval is €22,883. Viegas and Antunes (2013) classify these NUTS 3 (except for Barcelona, Girona, Lleida, and Tarragona) in their fourth interval. This interval encompasses incomes between €17,836.92 and €22,128.46, which is comparable to the average income observed in the sample under analysis. This allows for the examination of the growth in GDP per capita over the 2000-2019 period.

In conclusion, the provinces with the highest income levels are Álava, Madrid, Guipúzcoa, and Navarra, as determined in the fourth interval. The outcome is not unexpected, given Madrid's status as the capital of Spain and the fact that three of the provinces in question are situated in close proximity to France and benefit from special and highly advantageous regimes within the Spanish state. This classification is consistent with the NUTS 3 classification included in the fifth interval of Viegas and Antunes (2013), with incomes between €22,128.47 and €27,624.59, which are considerably lower than the values observed in our fourth level of the provincial classification.

Nevertheless, an examination of the Iberian Southwest sample reveals that Madrid is situated within the fourth interval, exhibiting the highest levels of GDP per capita. Lisbon and coastal Alentejo are positioned within the third interval, although their relatively low population density may contribute to this classification. However, most of the Iberian Southwest falls within intervals one and two, which correspond to the lowest levels of GDP per capita. Therefore, the counties of Beira Baixa, Cáceres, Médio Tajo, Badajoz, and Alto Alentejo are situated within the first interval, which encompasses the lowest levels of GDP per capita. The second interval comprises of Ciudad Real, Toledo, Central Alentejo and Leziria do Tajo, which exhibit low average levels of GDP per capita. These findings illustrate the economic fragility of the Iberian Southwest, apart from the capital cities of Madrid and Lisbon, as well as the port of Sines in coastal Alentejo (see table 4).

Table 4.
The average GDP (PPS) pc of the Iberian Peninsula. 2000-2019

NUTS 3	2000-2019	2000-2007	2008-2014	2015-2019
A Coruña	21345,00	18112,50	22485,71	24920,00
Lugo	20310,00	16800,00	21171,43	24720,00
Ourense	18935,00	15875,00	19585,71	22920,00
Pontevedra	20050,00	17512,50	20642,86	23280,00
Asturias	20980,00	18625,00	21814,29	23580,00
Cantabria	22095,00	20112,50	22585,71	24580,00
Álava	34750,00	30912,50	36485,71	38460,00
Gipuzkoa	30455,00	27112,50	31085,71	34920,00
Vizcaya	28595,00	24812,50	29571,43	33280,00
Navarra	29230,00	26675,00	29685,71	32680,00
La Rioja	25635,00	23287,50	26085,71	28760,00
Huesca	25810,00	22425,00	27171,43	29320,00
Teruel	24460,00	22225,00	25528,57	26540,00
Zaragoza	25705,00	23025,00	26142,86	29380,00
Madrid	32000,00	28437,50	32714,29	36700,00
Ávila	18535,00	16325,00	19200,00	21140,00
Burgos	26460,00	23462,50	27028,57	30460,00
León	19990,00	17837,50	20785,71	22320,00
Palencia	24040,00	20825,00	24742,86	28200,00
Salamanca	19290,00	17512,50	19485,71	21860,00
Segovia	22150,00	20987,50	22342,86	23740,00
Soria	23795,00	20937,50	24057,14	28000,00
Valladolid	24055,00	21500,00	24414,29	27640,00
Zamora	18310,00	15587,50	19471,43	21040,00
Albacete	18625,00	16225,00	19100,00	21800,00
Ciudad Real	19370,00	17225,00	19900,00	22060,00
Cuenca	19665,00	16700,00	20357,14	23440,00
Guadalajara	19575,00	18587,50	19785,71	20860,00
Toledo	18040,00	17100,00	18142,86	19400,00
Badajoz	16290,00	14050,00	16957,14	18940,00
Cáceres	16645,00	14112,50	16985,71	20220,00
Barcelona	28130,00	25450,00	28300,00	32180,00
Girona	27640,00	26137,50	27671,43	30000,00
Lleida	28195,00	25762,50	29100,00	30820,00
Tarragona	28405,00	26537,50	27857,14	32160,00
Alicante	18855,00	18312,50	18342,86	20440,00
Castellón	24620,00	23025,00	23585,71	28620,00
Valencia	22235,00	20412,50	22642,86	24580,00
Ibiza	22885,00	21075,00	22157,14	26800,00

TABLE 4. CONT.
The average GDP (PPS) pc of the Iberian Peninsula. 2000-2019

NUTS 3	2000-2019	2000-2007	2008-2014	2015-2019
Mallorca	20860,50	17359,99	21719,26	25259,05
Menorca	21260,00	21300,00	20514,29	22240,00
Almería	19785,00	19487,50	19142,86	21160,00
Cádiz	17345,00	16312,50	17514,29	18760,00
Córdoba	16695,00	14837,50	17171,43	19000,00
Granada	16860,00	14950,00	17300,00	19300,00
Huelva	18180,00	16862,50	18100,00	20400,00
Jaén	16370,00	14450,00	16728,57	18940,00
Málaga	17730,00	16450,00	17942,86	19480,00
Sevilla	19045,00	17200,00	19614,29	21200,00
Murcia	19745,00	18012,50	19971,43	22200,00
Ceuta	19860,00	18575,00	20242,86	21380,00
Melilla	18745,00	18175,00	18771,43	19620,00
El Hierro	17370,00	15237,50	19100,00	18360,00
Fuerteventura	22570,00	24200,00	20857,14	22360,00
Gran Canaria	19890,00	19325,00	19742,86	21000,00
La Gomera	18645,00	16762,50	19985,71	19780,00
La Palma	17865,00	15537,50	19214,29	19700,00
Lanzarote	23775,00	24225,00	22385,71	25000,00
Tenerife	21235,00	20412,50	21300,00	22460,00
Alto Minho	14150,00	11562,50	14714,29	17500,00
Cávado	15555,00	13187,50	15714,29	19120,00
Ave	15225,00	12975,00	15128,57	18960,00
Á. M. do Porto	18820,00	16862,50	18900,00	21840,00
Alto Tâmega	12140,00	9800,00	13100,00	14540,00
Tâmega e Sousa	11570,00	9700,00	11985,71	13980,00
Douro	13175,00	10487,50	13885,71	16480,00
T. de T. Montes	14475,00	11812,50	15428,57	17400,00
Algarve	20915,00	18825,00	20500,00	24840,00
Oeste	16315,00	15012,50	16185,71	18580,00
Região de Aveiro	19060,00	17150,00	18857,14	22400,00
Região de Coimbra	17965,00	15837,50	18300,00	20900,00
Região de Leiria	19550,00	17450,00	19742,86	22640,00
Viseu Dão Lafões	14680,00	12750,00	15100,00	17180,00
Beira Baixa	17150,00	14200,00	18071,43	20580,00
Médio Tejo	16400,00	14812,50	16628,57	18620,00
B. e S. da Estrela	12500,00	10400,00	12728,57	15540,00
Á.M. de Lisboa	27580,00	25475,00	28285,71	29960,00
Alentejo Litoral	24970,00	21825,00	24771,43	30280,00

Table 4. cont.
The average GDP (PPS) pc of the Iberian Peninsula. 2000-2019

NUTS 3	2000-2019	2000-2007	2008-2014	2015-2019
Baixo Alentejo	17805,00	14225,00	18871,43	22040,00
Lezíria do Tejo	17720,00	16650,00	17614,29	19580,00
Alto Alentejo	15150,00	13575,00	15100,00	17740,00
Alentejo Central	17775,00	16412,50	17542,86	20280,00
R. A. dos Açores	17535,00	15200,00	18228,57	20300,00
R. A. da Madeira	19295,00	17000,00	19814,29	22240,00
Total	**22883,19**	**20578,70**	**23277,39**	**26018,49**

Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

Table 5.
Average GDP (PPS) pc of the Southwest Iberian Peninsula. 2000-2019

NUTS 3	2000-2019	2000-2007	2008-2014	2015-2019
Madrid	32000,00	28437,50	32714,29	36700,00
Toledo	18040,00	17100,00	18142,86	19400,00
Cáceres	16645,00	14112,50	16985,71	20220,00
Badajoz	16290,00	14050,00	16957,14	18940,00
Ciudad Real	19370,00	17225,00	19900,00	22060,00
Beira Baixa	17150,00	14200,00	18071,43	20580,00
Medio Tajo	16400,00	14812,50	16628,57	18620,00
Leziria do Tajo	17720,00	16650,00	17614,29	19580,00
Alto Alentejo	15150,00	13575,00	15100,00	17740,00
Alentejo Central	17775,00	16412,50	17542,86	20280,00
Alentejo Litoral	24970,00	21825,00	24771,43	30280,00
A.M. Lisboa	27580,00	25475,00	28285,71	29960,00
Totals	**27212,98**	24325,49	27827,06	30973,25
Without capitals	**17607,84**	15757,92	17928,50	20118,80

Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

4.2. Dynamic variation and factorial decomposition of GDP (PPS) pc

A factorial decomposition of GDP (PPS) per capita has been conducted into its two constituent components. The utilisation of Gross Domestic Product (GDP) per employee (productivity) and employment per population (an approximation of the employment rate) as an analytical tool facilitates comprehension of the underlying factors that explicate the disparities in living standards between countries or regions. Furthermore, the cumulative average annual variation rates (CAAVR) for the three variables have been calculated for the subperiods of analysis, as well as for the entire period. The results are presented in condensed form for both the Iberian Peninsula and its NUTS 3 spatial units, and for the Iberian Southwest corridor, with and without the inclusion of capitals (see tables 6, 7, and 8).

Table 6 illustrates the categorisation of the various provinces in terms of their productivity and employment rates in 2019, representing the final year of the period under examination. The provinces have been classified according to both the productivity rate (vertical) and the employment rate (horizontal). In order to classify the productivity rate of the provinces, the value of each province has been taken and distributed into intervals in a homogeneous manner. The results led to the esta The results led to the

establishment of three intervals, namely low, medium, and high productivity, with twenty-eight units in each stratum. Moreover, the distribution of the employment/population ratio has resulted in the stratification of values into five categories: very low (0.03-0.28); low (0.28-0.42); medium (0.42-0.56); high (0.56-0.96); and very high (0.96-14.81).

It is evident that most provinces classified within the low productivity range are situated within the Southwestern Iberian Corridor. Nevertheless, their employment/population rates are distributed between the high and very high value ranges. Madrid is the sole province with high productivity, despite exhibiting low employment and population rates. Badajoz, Cáceres, and Ciudad Real are situated at an intermediate level of productivity yet display low employment and population values. It is evident that provinces with a very high employment rate also exhibit high productivity rates, with all Spanish provinces falling into this category. Conversely, provinces with low and very low employment rates are characterised by low productivity rates (see table 6).

TABLE 6.
GDP factorial decomposition (PPS): productivity/employment. 2000-2019

Employment/ Productivity	Very high	High	Medium	Low	Very low
Low		Albacete Cuenca Toledo Tarragona Castellón Menorca Almería	Málaga El Hierro Fuerteventura Gran Canaria A. Tâmega Douro Oeste R. Coimbra	R. de Leiria V.Lafões Beira Baixa Médio Tejo B.e S.Estrela	Á. M. Lisboa Alent. Litoral Baixo Alentejo Lezíria do Tejo Alto Alentejo Alent. Central R. A. Açores R. A. Madeira
Medium		Burgos Palencia Segovia Soria Valladolid Zamora Ciudad Real Guadalajara Badajoz Cáceres Barcelona Gerona Lérida Alicante Valencia Ibiza Mallorca	Córdoba Granada Jaén Ceuta Melilla La Palma Tenerife Alto Minho Cávado Ave T. e Sousa		

TABLE 6. CONT.
GDP factorial decomposition (PPS): productivity/employment. 2000-2019

Employment/ Productivity	Very high	High	Medium	Low	Very low
High	La Coruña Lugo Orense Pontevedra Asturias Cantabria Álava Gipuzkoa Vizcaya Navarra La Rioja Huesca Teruel Zaragoza Madrid	Ávila León Salamanca Cádiz	Huelva Sevilla Murcia La Gomera Lanzarote AM.Porto T. Montes Algarve R. de Aveiro		

Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

In examining the data presented in Table 7, it becomes evident that the primary driver of the cumulative average variations in GDP (PPS) per capita is productivity (GDP per employee). While this factor plays a significant role, the approximation to the employment rate (employment per population) also exerts a notable influence. Previous studies have indicated that disparities between regions are due to differences in productivity and employment rates (Rodríguez, 1999).

Moreover, it is noteworthy that during the period of economic crisis (2007-2014), the rates of variation in productivity remained positive, with figures of 1.78% for the Iberian Peninsula and 1.45% for the Iberian Southwest. In the absence of capital cities, the employment rate exhibits a contrasting trend, with negative values (-2.44 for the Iberian Peninsula and -2.67 and -1.88 for the Iberian Southwest with and without capitals, respectively). This result is consistent with the economic consequences of the global financial crisis (GFC) that began in 2008 and continued until 2012. During this period, Spain experienced a significant loss of employment, with more than 7 million jobs lost (Rocha & Aragón, 2012). Rocha and Aragón (2012) observed that the unemployment rate rose to over 8 million, resulting in a total of over 24 million unemployed individuals in Spain.

Furthermore, in the analysis of productivity, the average growth in the Iberian Peninsula has been 2.11%. For the sample of provinces in the South-West of the Iberian Peninsula with capitals, it has been 3.56%. In contrast, for the South-West of the Iberian Peninsula without capitals, it has been 1.95% for the entire period (see Table 7). Some authors posit that the evolution of productivity is the key to achieving convergence in Spain and Portugal (López Martínez, 2001).

Regarding the employment/population ratio, an approximation of the employment rate reveals that it has grown by a mere 0.18% over the entire period in the Iberian Peninsula. Conversely, it has decreased slightly in the South-West of the Iberian Peninsula without capitals (-0.05%) and has increased in the South-West of the Iberian Peninsula with capitals (0.30%) (see Table 7).

TABLE 7.
CAAVR (%) of GDP PPS pc: productivity/employment. 2000-2019

CAAVR PIB PPSpc				
	2000-2007	2007-2014	2014-2019	2000-2019
Iberian Peninsula	5,07	-0,7	3,44	2,48
Southwest Iberia with capitals	4,83	-0,46	3,35	2,46
Southwest Iberia without capitals	5,17	-0,94	3,78	2,52
CAAVR Productivity				
	2000-2007	2007-2014	2014-2019	2000-2019
Iberian Peninsula	3,62	1,78	0,5	2,11
Southwest Iberia with capitals	3,61	1,45	0,35	1,95
Southwest Iberia without capitals	4,12	1,8	1,5	2,57
CAAVR Employment				
	2000-2007	2007-2014	2014-2019	2000-2019
Iberian Peninsula	1,4	-2,44	2,21	0,18
Southwest Iberia with capitals	1,18	-1,88	2,17	0,3
Southwest Iberia without capitals	1,01	-2,69	2,25	-0,05

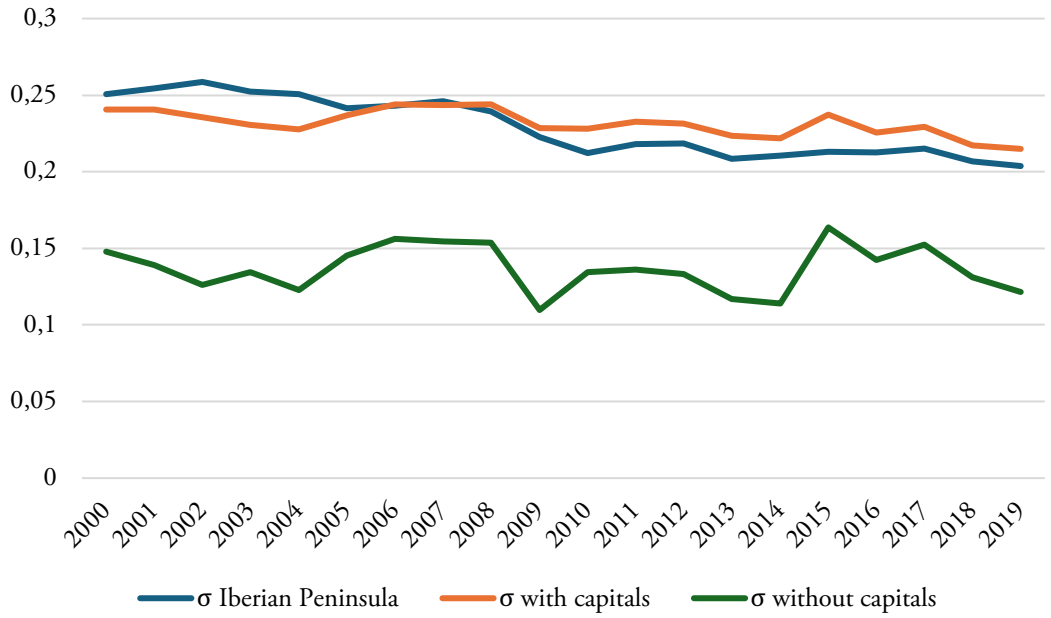
Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

4.3. σ CONVERGENCE

The evolution of σ convergence is illustrated in figure 1, which depicts both the Iberian Peninsula and the Iberian southwest corridor with and without capitals. The analysis of the period in question reveals a slight reduction in dispersion in both the Iberian Peninsula and in the Iberian Southwest Sorridor. With respect to the Iberian Peninsula, it can be observed that throughout the period from 2007 to 2010, a decrease in dispersion is evident, indicating the existence of economic convergence. From 2010 onwards, the trend line demonstrates fluctuations, yet economic convergence is evident throughout the analysed period.

However, in the case of the Iberian Southwest, excluding capital cities, dispersion analysis reveals a pattern of alternating convergence and divergence. Nevertheless, a slight convergence can be observed in the Iberian Peninsula from 2000 to 2019.

FIGURE 1.
σ Convergence of GDP (PPS) pc 2000-2019



Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

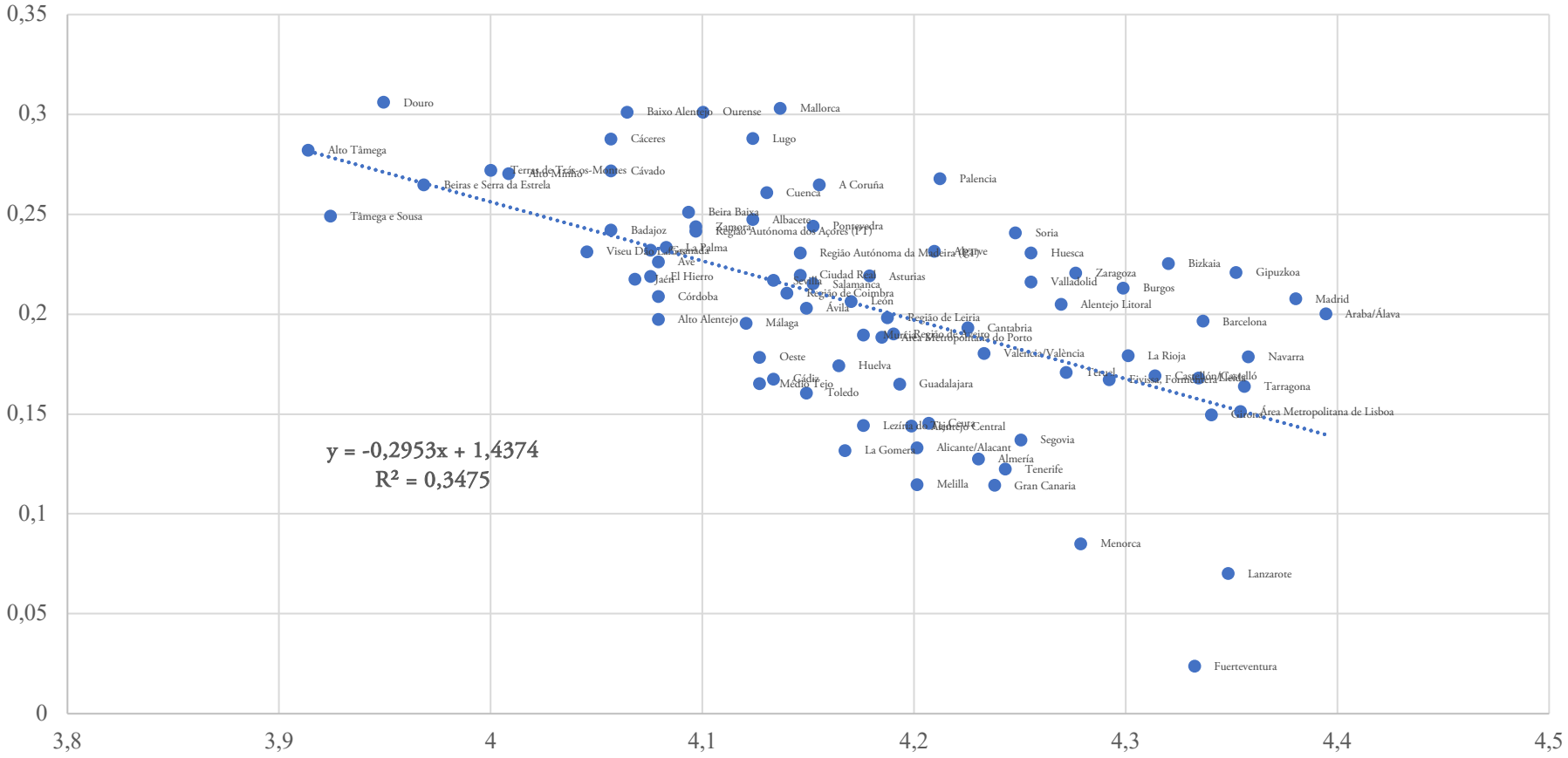
4.4. Conditional β convergence

Sala-i-Martín (1994) posits that β convergence is a necessary, though not sufficient, condition for σ convergence. Nevertheless, the calculation of β convergence is of interest as a complement to this convergence. Figure 2 illustrates the variation in terms of growth of the logarithm of GDP (PPS) pc compared to its situation in the starting year for the period 2000-2019.

The second chart presents the results of the beta-convergence analysis. The x-axis depicts the gross domestic product (GDP) per capita at the outset of the period, while the y-axis illustrates the growth rate of the provinces at the conclusion of the period. The regression line displays a negative trend, thereby substantiating the existence of beta-convergence. Furthermore, it is evident that the provinces which exhibited a higher GDP per capita at the outset of the period subsequently experienced the lowest growth. Conversely, the provinces that commenced the period with a lower GDP per capita demonstrated higher growth rates.

The Canary Islands commenced the period under examination with a relatively high GDP per capita yet exhibited a lower growth rate. Conversely, the provinces exhibiting the highest growth rates are predominantly Portuguese, representing regions with initially lower income levels, along with a few Spanish provinces, including Cáceres, Badajoz, Lugo, and Mallorca. Nevertheless, despite experiencing this growth, they remain included in the low or medium-low levels of GDP per capita. In his analysis of Spanish regions from 1985 to 1999, López (2001) identifies Madrid, the Balearic Islands and Catalonia as the leading performers, followed by Extremadura, and Andalucía. Our analysis reveals that the ranking of the regions has remained largely unchanged. The Islands, including the Canary Islands, Madrid, Lisbon, Navarre, and Álava, continue to lead the list, as do the two provinces that constitute Extremadura (Cáceres and Badajoz), along with certain Portuguese provinces.

FIGURE 2.
Conditional β convergence of the Iberian Peninsula. 2000-2019



Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

4.5. γ Convergence

The γ convergence has been calculated from the the Spearman's Rho γ rank concordance indicator for the GDP (PPS) per capita variable (see table 8). To establish whether there is convergence, we must examine the correlation result. If the correlation is positive and close to 1, it means that provinces that were relatively richer or poorer in 2000 remain in the same position. If, on the other hand, the correlation is low or negative, it suggests that the poorer provinces have grown faster, indicating the existence of convergence.

The results indicate that for the Iberian Peninsula (0.05965), the correlation is very low, indicating economic convergence between the provinces. That is, the provinces with the lowest GDP per capita in 2000 have grown faster than the high-income provinces.

On the other hand, for the analysis of the gamma-convergence of the provinces that make up the Iberian Southwest, we can say that, for the analysis with capitals, the ratio is low (0.293). Therefore, a low level of convergence can be observed. This is not the case for the analysis of the sample without capitals, where the result indicates that there is no convergence, as the relationship between the variables is moderate (0.515). This indicates that the provinces remain in the same position as in the year 2000.

Table 8.
γ Convergence of GDP (PPS). 2000-2019

Sample	Rho Spearman
Iberian Peninsula	0,059653741
Southwest Iberia with capitals	0,293706294
Southwest Iberia without capitals	0,515151515

Source: Own elaboration using data from the National Institute of Statistics (INE) in Spain and Portugal.

5. Conclusions

The principal aim of this paper is to examine the economic convergence between the various NUTS 3 regions of the Iberian Peninsula over the period 2000-2019. In order to determine whether economic convergence has been achieved in the Iberian Peninsula and the region known as the Iberian Southwest Corridor during this period, this study pursues the objective set by EU policy of removing cross-border barriers.

The analysis revealed that, in general, GDP (PPS) per capita has increased between the sub-periods under consideration, within the specified broad time frame, in both the Iberian Peninsula and the Iberian Southwest. In conclusion, it can be stated that the Iberian Peninsula is characterised by a high prevalence of income poverty, with 44 provinces exhibiting a below-average income level compared to the rest of the sample. The same provinces consistently rank among the poorest, namely Cáceres, Badajoz, Granada, Córdoba, and Jaén, which are located in the south and southwest of Spain. Conversely, the provinces with the highest income are Madrid and those situated in the north of Spain, which also exhibit the highest industrial activity. Furthermore, the economic weakness of the Iberian Southwest is evident, apart from the capitals, Madrid and Lisbon.

In general, the Spanish NUTS 3 regions have exhibited an upward trend throughout the period under analysis, not only in terms of GDP, but also in terms of population. In contrast, the Portuguese NUTS 3 regions have demonstrated a positive variation throughout the period under analysis in terms of income, but not in terms of population. This discrepancy is related to the serious internal demographic desertification affecting the country.

Considering the factorial decomposition of GDP per capita, it can be concluded that productivity and employment are fundamental variables for economic growth. It thus follows that the governments of

Spain and Portugal, including the autonomous governments, must base their public policies on the promotion of employment and the provision of support to companies to increase productivity.

The results of the convergence analysis indicate that economic convergence occurred in the Iberian Peninsula between 2000 and 2019. It is possible that this convergence is not a result of the increase in GDP per capita of the Spanish provinces, but rather a consequence of the economic crisis of 2008. At that time, the average income of the Iberian Peninsula was almost identical to the average income of the Iberian Southwest Corridor, suggesting that the Portuguese provinces were in a state of economic convergence with the Spanish provinces at that particular time. This leads to the conclusion that economic convergence exists.

Conversely, an analysis of beta-convergence also permits the conclusion that there is economic convergence. Although not the subject of our analysis, we identified the presence of clusters comprising spatial units with values that are more closely aligned. From the outset, we observed disparate areas exhibiting comparable growth rates and levels of GDP per capita.

The growth and convergence witnessed in Portugal and Spain towards the EU average are well documented. However, significant internal asymmetries persist, necessitating attention to achieve a more equitable and sustainable convergence.

As principal limitations, we propose the incorporation of supplementary variables to enhance this study, including an investigation of convergence in terms of wages, working conditions, and cross-border mobility. Additionally, we suggest the implementation of exploratory analyses that consider sectoral and territorial discrepancies between the various spatial units under examination. Conversely, the evident disparity in rates of variation, both in employment and productivity, is apparent both in space and time. This justifies the analysis of socioeconomic patterns to gain a deeper understanding of the underlying dynamics and to make more informed decisions.

In terms of future research, it is notable that, at the joint level of Spain and Portugal, there is a prevalence of independent axes of economic growth. The consolidation, definition, and optimisation of new axes of Iberian socioeconomic growth will provide greater articulation, harmonisation, and meaning to the joint progress of the Iberian Peninsula. Another future line of research aims to assess the impact of European policies (mainly EU structural and cohesion funds) on the process of economic convergence between Spain and Portugal. This will enable the identification of necessary adjustments to improve their effectiveness.

References

Barro, R. J., & Sala-i-Martin, X. (2009). *Economic growth*. Editorial Reverte.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy, 100*(2), 223-251.

Barro, R. J., & Sala-i-Martin, X. I. (1991). Convergence across states and countries. *Brookings Papers on Economic Activity*, 107-182.

Barro, R. J., Sala-i-Martin, X., Blanchard, O. J., & Hall, R. E. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 107-182.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics, 106*(2), 407-443.

Boyle, G. E., & McCarthy, T. G. (1999). Simple measures of convergence in per capita GDP: A note on some further international evidence. *Applied Economics Letters, 6*(6), 343-347.

Boyle, G. E., & McCarthy, T. G. (1997). A simple measure of beta convergence. *Oxford Bulletin of Economics and Statistics, 59*(2), 257-264.

European Commission. (2021, July 14). Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. EU cross-border regions: True laboratories of European integration. Publications Office of the European

Union. https://op.europa.eu/en/publication-detail/-/publication/1c01833b-e480-11eb-895a-01aa75ed71a1/language-en

Eurostat. (2023). Eurostat Regio, 30 June 2023. https://ec.europa.eu/eurostat/web/interactive-publications/regions-2023

Fernández, C. (2022). La Unión Europea, las redes transeuropeas y la necesidad de un corredor central. *Cuadernos de Estrategia*, (210), 9-30.

García, A. (2022). The corridors of the southwestern Iberian Peninsula. An unexpected European development network. *Torre de los Lujanes: Bulletin of the Royal Economic Society of Friends of the Country of Madrid, 78*, 235-252.

García Greciano, B., Raymond Bara, J. L., & Villaverde Castro, J. (1995). The convergence of the Spanish provinces. *Papers on Spanish Economy*, (64), 38-53.

García Salas, A. G., & Mora Aliseda, J. M. (2022). The Iberian southwest corridor. New European development axis. Infrastructure as a lever for economic and environmental sustainability. *Cuadernos de Estrategia, 210*, 169-182.

García-Solanes, J., Beyaert, A., & López-Gómez, L. (2021). Institutional convergence clubs in the euro area. *Spanish Economy Papers,* (168), 48-190.

Gouardères, F. (2024, March). La nomenclatura común de unidades territoriales estadísticas (NUTS). *Fichas temáticas sobre la Unión Europea*. Parlamento Europeo. https://www.europarl.europa.eu/factsheets/es/sheet/99/la-nomenclatura-comun-de-unidades-territoriales-estadisticas-nuts-

Hamadeh, N., Van Rompaey, C., & Metreau, E. (2023). World Bank Group country rankings by income level for fiscal year 24 (July 1, 2023 - June 30, 2024). *World Bank Blogs*. https://blogs.worldbank.org/en/opendata/world-bank-group-country-rankings-by-income-level-for-fiscal-year-24-july-1-2023-june-30-2024

INE (2021). Database of the Portuguese Statistics Institute. Retrieved 15 April, 2022. https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE

INE. (2021). The Iberian Peninsula in figures, 2020. https://www.ine.es/jaxiT3/Tabla.htm?t=2915&L=0

INE. (2019). Population by autonomous communities and cities and size of municipalities. Retrieved 15 April, 2022. https://www.ine.es/jaxiT3/Tabla.htm?t=2915&L=0

León, G. (2013). Economic growth and convergence: A review for Colombia. https://www.google.com/search?q=convergencia+economica+sala+i+martin+1990&oq=convergencia+economica+sala+i+martin+1990&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIHCAEQIRigAdIBCTEwNDQ3ajBqN6gCALACAA&sourceid=chrome&ie=UTF-8&safe=active&ssui=on

López Martínez, I. (2001). Nominal convergence and real convergence of Spain and Portugal with the European Union (1986-1999). *Applied Economics Studies, 19*(3), 49-68.

López-Villuendas, A. M., & del Campo, C. (2022). Analysis of economic convergence in the Galicia-Northern Portugal Euroregion in the period 1980-2019. *Galega de Economía Journal, 31*(2), 1-19.

Mora Corral, A. J. (2002). *On economic convergence. Theoretical aspects and empirical analysis for European and Spanish regions.* University of Barcelona.

Peres Lopes, L. (2004). A convergência da produtividade nas regiões NUTS III de Portugal continental o efeito da estrutura regional de emprego. *Revista Portuguesa de Estudios Regionais, 5*, 79-95.

Pinkovskiy, M., Sala-i-Martin, X., Chatterji-Len, K., & Nober, W. H. (2024). Inequality within countries is falling: Underreporting-robust estimates of world poverty, inequality and the global distribution of income. *National Bureau of Economic Research*. https://www.nber.org/papers/w32203

Rocha, F., & Aragón, J. (2012). The economic crisis and its effects on employment in Spain. *Gaceta Sindical, 19*(19), 67-90.

Rodríguez, M. M. (1999). Regional economic growth and convergence in Spain, in the long term. *Journal of Regional Studies, (54)*, 47-65.

Sala-i-Martin, X. (2012). The great convergence. *La Vanguardia*. Retrieved January 20, 2012.

Sala-i-Martin, X. X. (1996). Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review, 40*(6), 1325-1352.

Sala-i-Martín, J., Cuah, D. T., & Díaz Giménez, J. (1994). The wealth of the regions. Evidence and theories on regional growth and convergence. *Currency and Credit Journal, 198*, 13-80.

Torres, T., & Sala, M. (2008). The socio-economic convergence of the Spanish autonomous communities in the nineties. *Applied Economics Studies, 26*(3), 179-206.

Viegas, M., & Antunes, M. (2013). Convergence in the Spanish and Portuguese NUTS 3 regions: An exploratory spatial approach. *Intereconomics, 48*(1), 59-66.

Villaverde Castro, J., & Pérez González, P. (1996). The axes of growth in the Spanish economy. *Papers on Spanish Economy*, 63-78.

Villaverde, J. (2004). Provincial convergence in Spain: A spatial analysis. *Papers on Spanish Economy, 100*(1), 210-221.

ORCID

María Gema Flores Polán https://orcid.org/0000-0002-2031-7430

Gemma Fuentes Calle https://orcid.org/0000-0002-5415-895X

Luís Fernando De la Macorra Cano https://orcid.org/0000-0002-1828-4684

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 67-89
https://doi.org/10.38191/iirr-jorr.26.001

Articles

Does the location of company headquarters affect regional trade flows? Evidence from Spanish regions

Francisco Morillas-Jurado, Paz Rico-Belda**, Bernardí Cabrer-Borrás****

Received: 02 April 2024
Accepted: 09 December 2025

Abstract:

This paper examines trade flows between Spanish regions to assess whether the location of company headquarters affects regional trade flows. To achieve this, various gravity models are used to assess whether economic and geographical factors influence the intensity of trade flows. This study utilises European regional data, provided by the EUREGIO database, focusing specifically on Spanish regions. The results show that factors such as inertia, territorial contiguity, economic attraction between regions, and the location of company headquarters play a significant role in determining the intensity of GVA trade flows between Spanish regions.

Keywords: Trade flows; Input-Output table; gross value added; gravity model; spatial autocorrelation and distance.
JEL Classification: C67; O10; R50.

¿Influye la ubicación de la sede central de una empresa en los flujos comerciales regionales? Datos de las regiones españolas

Resumen:

En este trabajo se obtienen y analizan los flujos comerciales entre las regiones españolas con el objetivo de verificar si la localización de las sedes centrales de las empresas influye en la intensidad de dichos flujos. Para ello, mediante modelos de gravedad, se evalúa si factores económicos y geográficos influyen en la intensidad de los flujos comerciales. Este estudio utiliza datos regionales europeos, proporcionados por la base de datos EUREGIO, y se centra específicamente en las regiones españolas. Los resultados muestran que factores como la inercia, la contigüidad territorial, la atracción economica entre regiones y el efecto de la localización de la sede central de las empresas influyen en la intensidad de los flujos comerciales del VAB entre las regiones españolas.

Palabras clave: Flujos comerciales; tabla de insumo-producto; valor añadido bruto; modelo de gravedad; autocorrelación espacial y distancia.
Clasificación JEL: C67; O10; R50.

* Departament d'Economia Aplicada, Facultat d'Economia, Universitat de València, Spain. Francisco.Morillas@uv.es
** Departament d'Análisis Econòmica, Facultat d'Economia, Universitat de València, Spain. Paz.Rico@uv.es
*** Departament d'Análisis Econòmica, Facultat d'Economia, Universitat de València, Spain. Bernardi.Cabrer@uv.es
Corresponding author: Paz.Rico@uv.es

1. Introduction

The momentum observed in trade flows of goods and services, driven by globalisation and production relocation, has both positive and negative repercussions for the economies of the regions. The development of trade flows offers great opportunities that lead to economic growth (Frankel and Romer, 1999; Winters et al., 2004). However, knowing the mechanisms of their development is important in order to be able to intervene, where necessary, and achieve economic and/or social improvements. Given that the production and distribution of goods and services directly impact the location and/or relocation of production centres, these also indirectly affect the territorial distribution of income between the different local Economic Activity Units (EAU) of the companies. Better understanding of the interrelationship of interregional flows would allow the design and implementation of economic policies aimed at mitigating potential interregional imbalances.

Reuveny and Li (2003) evaluate the effects of trade in terms of income imbalances and conclude that trade reduces this imbalance. Thissen et al. (2019) consider that economic development is interregional by nature, with the regional dimension of economic development being increasingly at the centre of political and academic debates. These authors emphasise that economic growth is determined by physical and technological proximity, which is facilitated by interregional and national cross-border interactions in trade, investment and knowledge. Similarly, economic disparities can generate social unrest and feelings of comparative grievance between regions (He and Duchin, 2009), which can be the source of conflicts. Therefore, knowing how interregional trade flows develop and evolve is crucial. A better knowledge of interregional trade flows will aid in formulating more appropriate economic policy decisions.

This paper analyses the gross value added (GVA) flows of Spanish regions to examine how the location of company headquarters affects these flows. We propose a methodology to obtain and analyse regional GVA flows, focusing on interregional flow matrices derived from the equilibrium models of Leontief (1953) and Ghosh (1958). The application of this methodology focuses on quantifying trade flows between Spanish regions, using the EUREGIO database which contains annual data for the period 2000-2010 for European regions (see Thissen et al., 2018). From this database, the GVA flows are derived, allowing for a characterisation of regions based on their dependence, orientation and intensity of flows. Once the GVA flows are obtained, the analysis proceeds using two complementary approaches.

The first approach carries out a descriptive study of the GVA flows. This provides a global idea of the interdependence between the regions, evaluating the exchanges between two regions: region *r*, of origin, and region *s*, of destination.

The second approach analyses the flows using a gravity model that incorporates geographical location factors, in line with the work of Santamaría et al. (2023). In economic literature, gravity models are widely used to explain trade flows between regions (Evenett and Keller, 2002). These models typically use (i) trade flows between regions as the endogenous variable, and (ii) explanatory variables characterising the asymmetric distribution of trade flows. Thus, some regressors take into account characteristics such as inertia, distance, contiguity, and income. The sample period covers the years 2000-2010, resulting in a sample of 3,179 observations.

This paper seeks to contribute to the current literature in several ways. Firstly, in this study, the accounting method is considered the most appropriate method for obtaining the regional GVA flow matrix, since this method adequately accounts for the GVA imputed to each region when company headquarters are located outside the production centres. Secondly, a gravity model is used to assess whether economic and geographic factors affect interregional flows, emphasising the effect of the location of company headquarters.

Following this introduction, the paper is structured as follows. Section 2 presents the theoretical framework. Section 3 includes the description and analysis of the database used. Section 4 presents the gravity model. Section 5 shows the empirical application of the gravity model to Spanish regions. The final section offers the conclusions of the study.

2. Theoretical framework

Traditional gravity models (Evenett and Keller, 2002; Baldwin and Taglioni, 2006; Anderson, 2010; Head and Mayer, 2014; Head and Mayer, 2021) are rooted in an adaptation of Newton's law of gravitation, which relates to the force of attraction between two objects, depending on their masses and the distance separating them. Similarly, the gravity model is a tool used in a wide range of empirical fields (Baldwin and Taglioni, 2006; Kabir et al., 2017) to measure trade between countries. The Tinbergen (1962) model, one of the earliest macroeconomic models, suggests that trade between two countries depends mainly on two factors: the economic size of the countries and the distance between them. Building on this and inspired by the spatial econometrics proposed by Anselin (1988, 2003), an additional factor derived from substantive spatial autocorrelation or dependence is incorporated.

The gravity model has played an important role in research on international trade, as it provides a simple and empirical way to evaluate and predict trade flows between countries. While Tinbergen's original model was relatively simple, it has since been expanded and modified to include additional factors that affect trade, such as trade policies, tariff barriers, and the characteristics of the economies. Head and Mayer (2014) provide a comprehensive review of the literature that uses the gravity model to explain trade flows between countries. More recently, Kabir et al. (2017) and Shahriar et al. (2019) have conducted a thorough review of recent developments in the gravity model applied to trade flow analysis. The gravity model remains a widely used tool in contemporary research. For instance, Nitsch and Wolf (2013) demonstrate the persistence of the border effect in internal German trade following reunification. Similarly, Wrona (2018) uses the gravity model to measure intranational trade integration in Japan and examine the existence of a border effect. Mayer et al. (2019) evaluate the benefits of intra-European trade promoted by the economic integration of the European Union, as well as the potential impact of its disintegration. Serrano-Domingo et al. (2020) analyse the immigration flows between countries. Yotov (2021) argues that trade flows between regions within the same country have been insufficiently considered and analyses the importance of domestic trade flows in the estimation of the gravity model of trade. Finally, Santamaría et al. (2023) analyse regional trade patterns in Europe, where distance, borders, and political divisions are key determinants.

This paper specifies a gravity model, following the approach of Santamaría et al. (2023), which includes the effect of the location of company headquarters. Head and Mayer (2019) refine traditional gravity models by including a new type of friction, highlighting the challenges of selling in markets geographically distant from a company's headquarters. Similarly, Wang (2021) examines how the location of corporate headquarters affects international trade dynamics. Using Chinese customs data, Wang finds that the headquarters effect accounts for approximately 20% of China's exports in 2000, and he argues that ignoring this effect could introduce bias into quantitative analyses of trade between the USA and China.

3. Database and descriptive analysis

This study uses the EUREGIO database, see Thissen et al. (2018), which provides information on interregional flows between the NUTS II regions of the EU as well as a supra-regional entity labelled "rest of the countries of the world", encompassing a total of 248 regions. The database covers the period from 2000 to 2010 and includes data on 14 production sectors.

To analyse the Spanish regional economy the information provided by EUREGIO must first be reorganised through successive aggregations. Specifically, the 248 regions are grouped into 21 and all 14 sectors are aggregated into a single sector. The 21 regions include the seventeen Spanish regions, together with the two Spanish autonomous cities Ceuta and Melilla, plus two other supra-regions; the first supra-region encompasses the rest of the EU's NUTS II, and the second supra-region includes the rest of the countries in the world that the database includes. This gives the matrices of interregional flows of intermediate consumption of goods and services (Matrix $\mathbf{Z}$) and the matrix of interregional flows of final demand for goods and services (Matrix $\mathbf{F}$) for the 21 regions (see Tormo et al., 2023).

The GVA matrix is not generally accessible and must be estimated. For its estimation, the EUREGIO database provides the flows of intermediate goods and services matrix, **Z**, and the flows of final goods and services matrix, **F**. Using the accounting method, based on the Leontief (1953) model, the production flow matrix **X** is obtained as the sum of **Z** and **F**. Subsequently, the GVA flow matrix, **V**, is calculated as the difference between $\mathbf{X}'$ and **Z**.

The multi-regional input-output Leontief model is expressed by the following equation[1]:

$$z_{ij}^{rs} + f_{ij}^{rs} = x_{ij}^{rs} \tag{1}$$

where the superscripts denote regions *r* and *s*, and subscripts denote sectors *i* and *j*.

This relationship can be expressed in matrix terms by the equation **(1.a)**:

$$\mathbf{Z} + \mathbf{F} = \mathbf{X} \tag{1.a}$$

where[2]:

Z is a square matrix of interregional flows of intermediate consumption $\mathbf{Z}=[z^{rs}]$.

F is a square matrix of interregional flows of demand for final goods $\mathbf{F} = [f^{rs}]$

X is a square matrix of interregional flows of production $\mathbf{X} = [x^{rs}]$

Similarly, from the supply perspective, the Ghosh (1958) model is used to express by the equation **(2)**:

$$z^{rs} + v^{rs} = x^{rs} \tag{2}$$

This relationship can be expressed in matrix terms as the equation **(2.a)**:

$$\mathbf{Z} + \mathbf{V} = \mathbf{X}' \tag{2.a}$$

where **V** is a square matrix of interregional flows of value added $\mathbf{V}=[v^{rs}]$.

From the equation **(1.a)** the matrix of production **X** is obtained. Similarly, from the equation **(2.a)** the GVA[3] flow matrix is obtained, matrix **V**.

In summary, the process described provides four matrices of interregional flows for each year of the period 2000 to 2010: a total of 44 square matrices of interregional flows of dimension 21 x 21[4].

The matrices of interregional flows enable the characterisation of regions based on the intensity of the interregional flows. Intensity is defined as the volume of flows originating in region *r* and destined for one or more other regions.

With respect to the results obtained for the GVA flow matrix using the accounting method, the interpretation of an Input-Output Table from the supply perspective is carried out by studying, by columns, the flow matrices **Z**, **V** and $\mathbf{X}'$. The concept is as follows: to produce a unit of a good or service requires primary inputs from the region itself and from other regions. These inputs configure an interregional square matrix of primary input flows, whose generic element is z^{rs}. The production of these goods and services in a region may originate in the same region or in other regions. The flows of the goods and services produced configure the square matrix of interregional production flows, whose generic element is x^{rs}. The difference between the matrix of interregional production flows and the matrix of interregional flows of primary inputs defines the matrix of GVA flows, **V**, whose generic element is v^{rs}.

[1] The standard input-output notation is used, where bold capital letters denote matrices (X) and bold lowercase letters denote vectors (x).

[2] In this paper, sectors are aggregated into a single sector, so the analysis focuses exclusively on regions.

[3] There are different techniques to estimate GVA. Among these are the expenditure method, the income method, and the value-added method (see Appendix A).

[4] Data can be provided by the authors upon request.

This process of deriving the GVA flow matrix from the intermediate consumption flow matrix and the production matrix is referred to as the accounting method.

It is important to emphasise that some elements of the GVA flow matrix, **V**, have a negative sign[5]. The negative sign in a GVA flow would indicate, for a given region *r*, that the flow of production of goods and services, x^{rs}, coming from this region *r* is lower than the flow of primary inputs z^{rs}, from the same origin. The negative value of an interregional flow of the GVA v^{rs} (when analysing the matrix **V**, interregional purchases, by columns) would be to the detriment of the purchasing region *r* (since its GVA ultimately decreases) and to the benefit of the selling region *s*. In this case, using the accounting method, an underestimation of the GVA of region *r* would occur.

From an economic point of view, GVA is an economic macromagnitude obtained from the difference between production and primary inputs. Furthermore, the GVA is basically configured by the sum of the remuneration of the capital factor, the gross operating surplus, and the remuneration of the labour factor.

Table 1 presents the GVA trade flow matrix for the year 2009 obtained using the accounting method using equation **(2.a)**. Analysing the matrix by rows and focusing on negative values reveals the GVA received from other regions. The regions that benefit most from interregional GVA flows are the following: Community of Madrid, Catalonia, and the Basque Country. The sum of the negative elements by column indicates the amount of GVA that each region gives to the others. The most disadvantaged in percentage terms are Extremadura, the Region of Murcia and the Canary Islands.

Figure 1 confirms that the regions that benefit the most are the Community of Madrid, followed by Catalonia and the Basque Country, regions where the largest number of company headquarters are concentrated. The regions most negatively affected by interregional GVA flows are Extremadura, the Region of Murcia and the Canary Islands, followed by Castilla-La Mancha, La Rioja, Castilla-León, the Balearic Islands, and Cantabria.

In conclusion, the presence of both positive and negative interregional flows of GVA between regions implies that some regions benefit from trade flows, while others are disadvantaged. Furthermore, it is important to note that negative GVA flows entail, from an accounting point of view, an underestimation (overvaluation) of the regional GVA, and consequently an underestimation (overvaluation) of its own GDP[6].

These results invite further study of the possible causes of these imbalances, which include, among others, the *effect of territorial contiguity*, the company *headquarters effect*, *temporal inertia,* and *spatial inertia*.

To measure the behaviour of a region through its trade flows, the statistical coefficient, CP^{rs} (Pavía-Miralles and Cabrer-Borrás, 2004), is used. The coefficient CP^{rs} is expressed as:

$$CP^{rs} = \frac{1}{2}\left(\frac{x^{rs}}{\sum_s x^{rs}} + \frac{x^{rs}}{\sum_r x^{rs}}\right)$$

This coefficient distinguishes the links between regions according to the origin of each trade flow and quantifies the relative importance of interregional flows by weighting each trade flow by the total flows of the regions. The CP^{rs} coefficients highlight, within the total set of flows, those flows that are of special importance for the sending region, the receiving region or both.

[5] The possibility that the GVA flow in a production sector (region) presents a negative value is considered in: EUROSTAT (1995); Hernández García (1997) and the Official Journal of the European Union (2013).

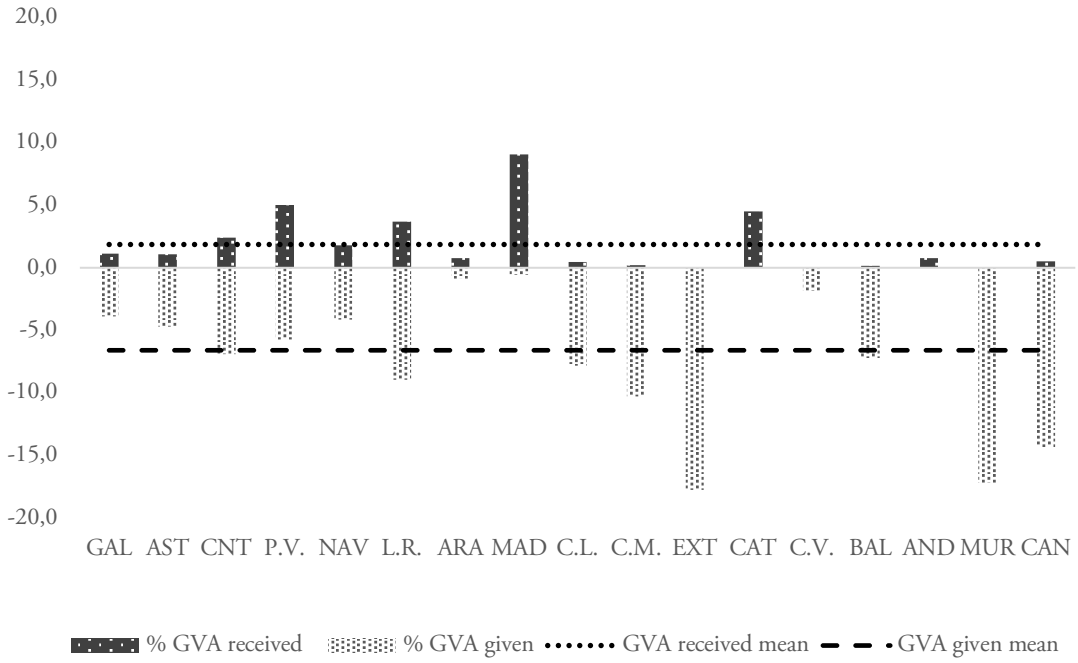
[6] The GDP is obtained after adding to the GVA the indirect taxes that are levied on production and subtracting the subsidies that companies receive.

TABLE 1.
Matrix of GVA interregional flows. Year 2009. (Units million €)

	GAL	AST	CNT	P.V.	NAV	L.R.	ARA	MAD	C.L.	C.M.	EXT	CAT	C.V.	BAL	AND	MUR	CAN
GAL	50 971	*-41*	*-160*	976	*-74*	*-175*	266	5 141	220	*-88*	296	2 577	752	*-56*	1 373	279	772
AST	352	20 615	5	462	63	*-92*	319	2 318	218	122	334	1 305	594	*-32*	1 001	*-111*	189
CNT	328	85	10 355	387	55	*-33*	163	1 030	260	155	*-53*	936	500	99	691	*-197*	323
P.V.	*-71*	*-68*	*-122*	54 103	*-272*	14	*-251*	6 137	*-446*	*-579*	*-546*	2 699	654	*-226*	978	*-627*	306
NAV	350	42	*-7*	807	13 986	70	330	1 258	200	88	*-109*	1 337	552	51	782	*-163*	354
L.R.	272	134	61	280	163	5 876	297	285	214	151	18	450	337	71	264	*-255*	135
ARA	371	*-20*	*-87*	807	*-43*	*-61*	26 507	1 761	157	-39	33	2 458	874	223	1 052	23	362
MAD	*-1 925*	*-817*	*-84*	*-3 218*	19	138	-34	131 581	-3 043	-1 996	-866	5 267	*-1 307*	*-640*	1 246	*-1 594*	*-1 222*
C.L.	717	281	85	1 661	183	*-44*	598	6 134	51 680	306	641	2 727	1 456	*-151*	1 581	819	-25
C.M.	767	209	35	1 481	213	*-41*	792	5 297	599	32 836	742	2 916	1 653	205	1 536	828	439
EXT	753	554	885	1 573	966	626	1 332	3 689	730	1 130	15 130	3 752	1 548	1 255	2 461	403	1 119
CAT	*-161*	*-207*	*-146*	*-455*	*-274*	*-34*	*-507*	*-1 030*	*-606*	*-612*	*-927*	123 276	*-422*	*-601*	1 674	*-1 145*	-204
C.V.	808	172	77	1 172	310	57	749	6 397	148	191	183	3 987	84 377	539	2 170	1 404	1 225
BAL	601	183	*-13*	669	130	*-30*	359	2 383	572	289	78	2 220	781	20 604	1 265	348	436
AND	*77*	*-101*	*-100*	1 098	46	*-30*	289	4 247	82	*-139*	*-297*	1 599	479	850	100 109	*-282*	19
MUR	1 059	1 081	1 069	1 777	1 148	756	1 605	4 936	394	1 349	977	4 751	1 680	1 350	2 987	25 182	1 548
CAN	475	210	*-94*	615	-27	*-79*	277	3 545	938	362	127	1 851	731	93	1 490	500	31 784
Total GVA	55 794	22 326	11 773	64 243	16 628	6 934	33 134	185 313	52 370	33 586	15 787	164 287	95 340	23 698	122 738	25 452	37 604
% GVA	5.77%	2.31%	1.22%	6.64%	1.72%	0.72%	3.43%	19.16%	5.42%	3.47%	1.63%	16.99%	9.86%	2.45%	12.69%	2.63%	3.89%
% GVA given	-3.87%	-4.69%	-6.91%	-5.72%	-4.14%	-8.93%	-0.86%	-0.56%	-7.82%	-10.28%	-17.72%	0.00%	-1.81%	-7.20%	0.00%	-17.18%	-14.31%
% GVA received	1.11%	1.08%	2.41%	5.02%	1.77%	3.68%	0.79%	9.05%	0.47%	0.18%	0.00%	4.49%	0.00%	0.18%	0.77%	0.00%	0.00%
Production	113 431	42 930	21 190	121 891	30 849	10 995	58 129	227 697	105 927	65 811	26 044	144 325	160 703	41 439	120 289	48 907	1 699
% Prod.	8.05%	3.05%	1.50%	8.65%	2.19%	0.78%	4.12%	16.15%	7.51%	4.67%	1.85%	10.24%	11.40%	2.94%	8.53%	3.47%	4.67%
% GVA / Prod.	37.30%	39.10%	40.60%	37.10%	38.00%	41.10%	38.90%	48.30%	39.30%	39.70%	39.40%	79.60%	44.60%	43.80%	77.60%	34.50%	44.30%

Note: GAL= Galicia, AST= Asturias; CNT= Cantabria: P.V.= Basque Country; NAV= Navarre; L.R.= La Rioja; ARA= Aragon; MAD = Community of Madrid; C.L.= Castilla-León; C.M .= Castilla-La Mancha; EXT = Extremadura; CAT = Catalonia; C.V.= Valencian Community; BAL= Balearic Islands; AND= Andalusia; MUR= Region of Murcia; CAN= Canary Islands.
Source: compiled by the authors based on the matrices of GVA interregional flows for 2009, EUREGIO.

FIGURE 1.
Regions benefitting from/disadvantaged by % GVA flows



Note: GAL= Galicia, AST= Asturias; CNT= Cantabria: P.V.= Basque Country; NAV= Navarre; L.R.= La Rioja; ARA= Aragon; MAD= Community of Madrid; C.L.= Castilla-León; C.M.= Castilla-La Mancha; EXT= Extremadura; CAT= Catalonia; C.V.= Valencian Community; BAL= Balearic Islands; AND= Andalusia; MUR= Region of Murcia; CAN= Canary Islands.
Source: compiled by the authors using data from Table 1.

Table 2 presents the estimates of the CP^{rs} coefficients for the year 2009, which have been calculated from the GVA trade flow matrix. The following results in Table 2 are of particular interest:

- The **CP** matrix presents negative values for some of the coefficients, CP^{rs}, affirming that they truly reproduce the direction of the "original" trade flows.
- The elements of greatest magnitude are found on the main diagonal of the matrix, which indicates that the most relevant GVA flows correspond to the internal flows of each region, produced within each of the regions.
- The analysis by columns of the **CP** matrix allows us to verify that there are several coefficients that stand out from the rest; those corresponding to the autonomous communities of Madrid and Catalonia. These results indicate that these are regions with a strong capacity to emit GVA flows.
- When analysing the elements of the **CP** matrix by row, the most important values are seen to be concentrated in the regions of Extremadura, Murcia, and Castilla-La Mancha, which indicates their high receiving capacity for GVA flows.

Figure 2 shows the main GVA flows between the different regions. The maps show that the most intense flows correspond to the regions closest to each other, and that the most relevant relative flows originate in the regions of Madrid and Catalonia. Therefore, these results corroborate the fact that the largest regions are those with greater intensity in trade flows.

TABLE 2.
Matrix of CP* calculated from the matrix of flows of GVA (Matrix V)

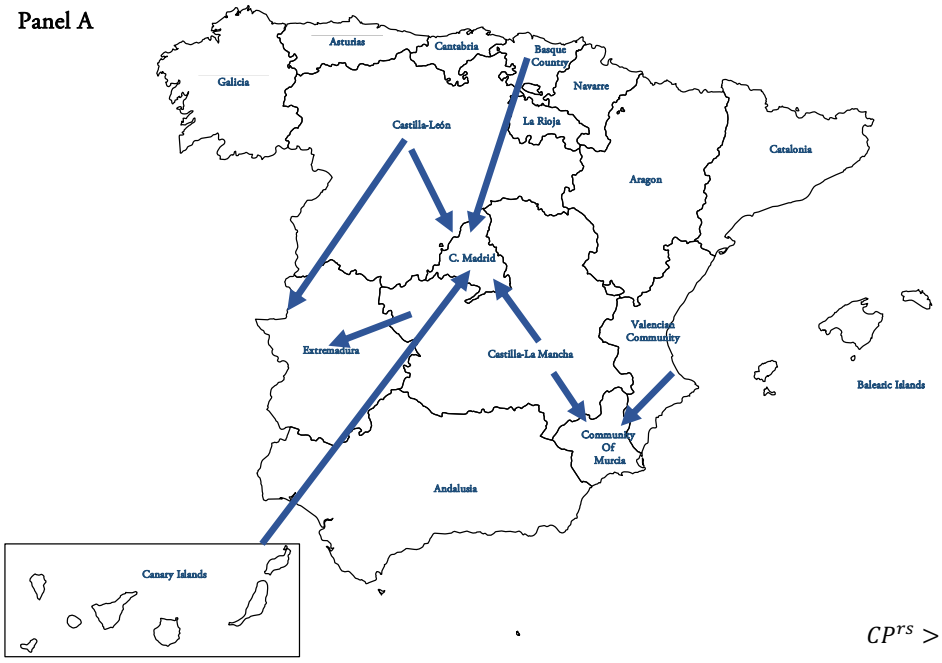
	GAL	AST	CNT	P.V.	NAV	L.R.	ARA	MAD	C.L.	C.M.	EXT	CAT	C.V.	BAL	AND	MUR	CEU	MEL	CAN
GAL	0.820	0.001	-0.005	0.014	0.000	-0.009	0.007	0.036	0.002	-0.002	0.032	0.020	0.009	0.002	0.012	0.027	-0.004	-0.001	0.018
AST	0.007	0.769	0.001	0.009	0.003	-0.008	0.010	0.028	0.004	0.002	0.025	0.019	0.009	0.000	0.017	-0.001	-0.001	0.031	0.008
CNT	0.009	0.005	0.694	0.010	0.003	-0.005	0.007	0.012	0.008	0.004	0.001	0.023	0.011	0.005	0.019	-0.005	-0.001	0.013	0.014
P.V.	-0.002	0.001	0.000	0.799	-0.001	0.011	-0.001	0.047	-0.008	-0.006	-0.006	0.019	0.009	-0.001	0.007	0.005	-0.008	-0.008	0.011
NAV	0.008	0.003	0.000	0.018	0.688	0.005	0.012	0.017	0.004	0.003	-0.001	0.025	0.012	0.003	0.017	0.000	-0.002	-0.002	0.013
L.R.	0.013	0.010	0.007	0.008	0.014	0.653	0.016	-0.001	0.010	0.007	0.006	0.020	0.012	0.006	0.013	-0.008	0.007	0.005	0.009
ARA	0.005	0.000	-0.003	0.012	0.000	-0.001	0.720	0.013	0.001	-0.001	0.015	0.029	0.012	0.010	0.013	0.017	-0.004	0.001	0.011
MAD	-0.013	-0.002	0.023	-0.024	0.016	0.033	0.014	0.812	-0.027	-0.012	0.014	0.020	0.000	-0.004	0.007	0.010	-0.011	0.016	-0.015
C.L.	0.011	0.009	0.006	0.025	0.008	0.002	0.015	0.047	0.820	0.007	0.043	0.020	0.018	0.000	0.015	0.035	-0.004	-0.005	0.004
C.M.	0.012	0.009	0.005	0.020	0.009	0.001	0.019	0.045	0.009	0.738	0.050	0.029	0.021	0.010	0.017	0.043	-0.006	0.006	0.012
EXT	-0.004	0.004	0.026	0.017	0.024	0.025	0.015	0.023	-0.005	0.002	0.464	0.037	0.007	0.022	0.021	0.008	0.039	0.048	0.016
CAT	0.000	0.002	0.003	0.001	0.001	0.006	-0.003	-0.001	-0.003	-0.003	-0.005	0.817	0.002	-0.009	0.012	0.011	-0.005	0.027	0.000
C.V.	0.009	0.011	0.014	0.012	0.016	0.015	0.019	0.032	0.000	0.006	0.030	0.021	0.782	0.014	0.014	0.064	-0.004	-0.005	0.021
BAL	0.012	0.008	0.000	0.013	0.007	-0.001	0.012	0.034	0.012	0.007	0.006	0.036	0.016	0.707	0.022	0.024	0.004	0.005	0.015
AND	0.003	0.002	0.001	0.018	0.006	0.005	0.010	0.023	0.001	0.002	0.016	0.011	0.009	0.024	0.796	0.018	0.105	-0.008	0.003
MUR	-0.003	0.020	0.034	0.009	0.027	0.037	0.014	0.022	-0.011	0.003	0.036	0.032	-0.007	0.017	0.019	0.515	0.043	0.002	0.012
CEU	0.006	0.003	0.004	0.009	0.005	-0.002	0.006	0.029	0.005	0.008	0.004	0.023	0.012	0.006	0.003	0.004	0.549	0.002	0.003
MEL	0.006	0.001	0.004	0.010	0.006	0.001	0.007	0.017	0.008	0.008	0.002	0.021	0.011	0.009	0.013	0.003	0.006	0.578	0.006
CAN	0.007	0.008	-0.003	0.009	0.000	-0.005	0.007	0.042	0.016	0.008	0.007	0.023	0.013	0.004	0.021	0.027	0.001	0.001	0.731

Note: GAL= Galicia, AST= Asturias; CNT= Cantabria: P.V.= Basque Country; NAV= Navarre; L.R.= La Rioja; ARA = Aragon; MAD = Community of Madrid; C.L.= Castilla-León; C.M. = Castilla-La Mancha; EXT= Extremadura; CAT= Catalonia; C.V.= Valencian Community; BAL= Balearic Islands; AND= Andalusia; MUR= Region of Murcia; CEU= Autonomous City of Ceuta; MEL= Auton. City of Melilla; CAN= Canary Islands.
Source: compiled by the authors based on the EUREGIO database.

* It should be noted that the CP^{rs} coefficients presented are derived from Pavía-Miralles and Cabrer-Borrás (2004). Specifically, they represent the average of the coefficients from the GVA flow matrix and its transposed matrix, thereby capturing the intensity of flows both in terms of GVA acquisition and transfer.

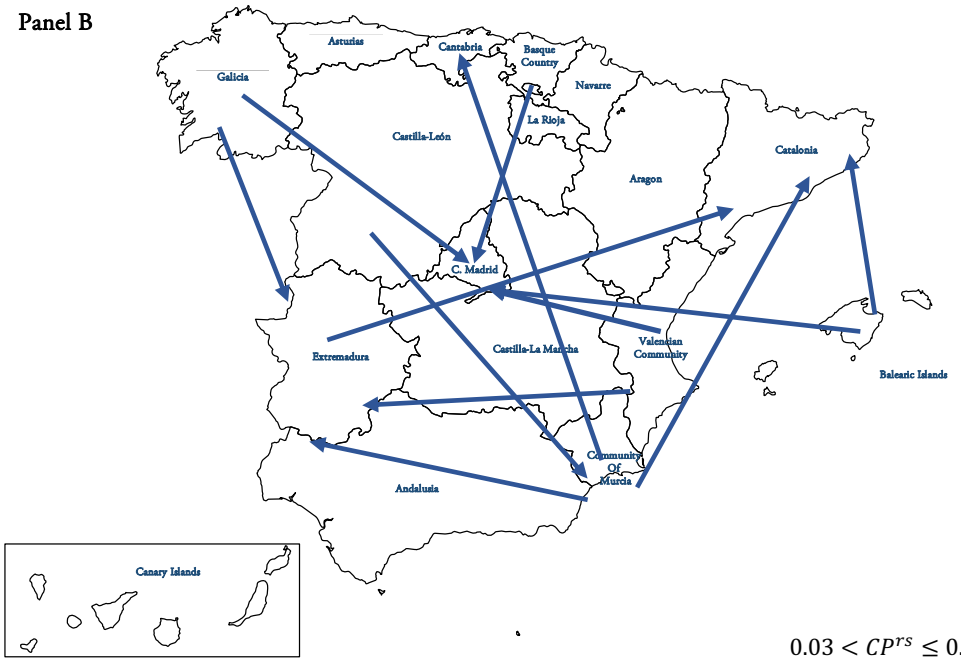
FIGURE 2.
GVA flows between regions using the coefficient CP^{rs}

Panel A



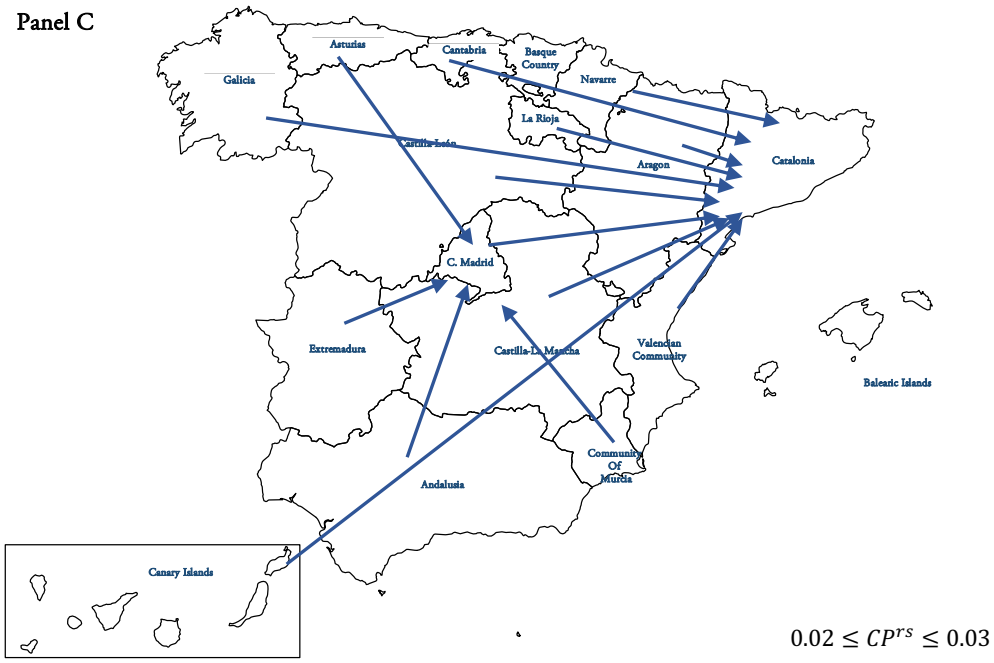
$CP^{rs} > 0.04$

Panel B



$0.03 < CP^{rs} \leq 0.04$

FIGURE 2. CONT.
GVA flows between regions using the coefficient CP^{rs}

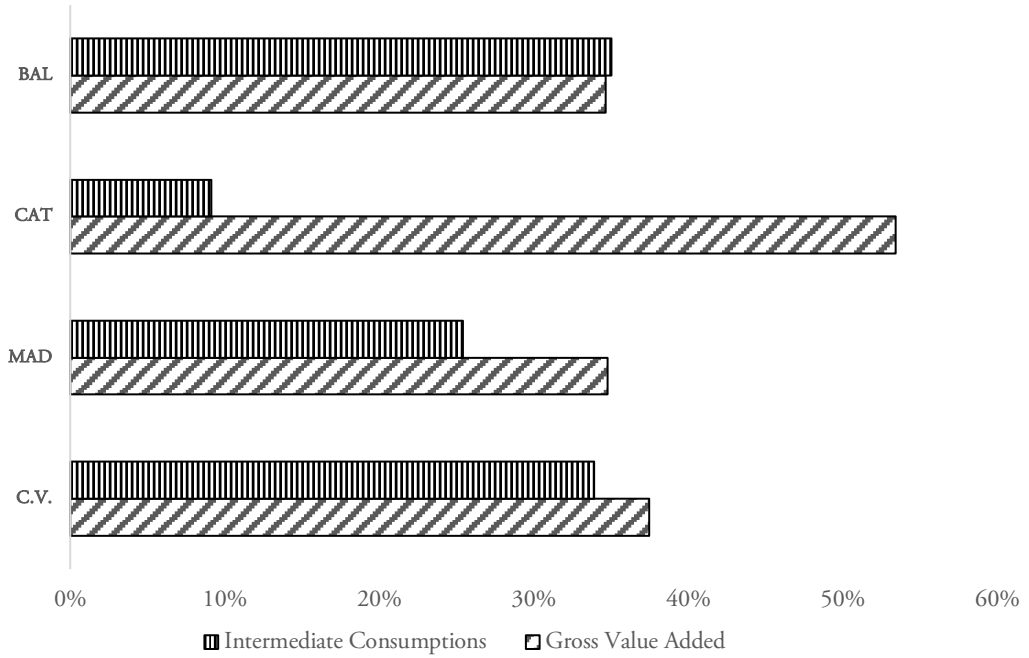


Note: The values of intraregional trade flows (the elements of the main diagonal of matrix V) have not been considered in the analysis. The arrows represent the flow of GVA between regions using the coefficients, in **CP** matrix, that exceeds the specified thresholds. Their direction indicates the origin region and the receiving region. Additionally, the thickness of the arrows represents the intensity of the flows. Panel C shows the information only for Madrid and Catalonian regions.
Source: compiled by the authors based on Table 2.

Figure 3 shows the production structure of the regions of the Balearic Islands, Catalonia, Community of Madrid and the Valencian Community, which are analysed through the intraregional flows of intermediate consumption, matrix **Z**, and GVA, matrix **V**. The results show that the production structures, from a supply perspective, are similar for the Community of Madrid, the Valencian Community and the Balearic Islands, while the Catalonia production system is characterised by presenting very high values in GVA flows to the detriment of intermediate consumption flows.

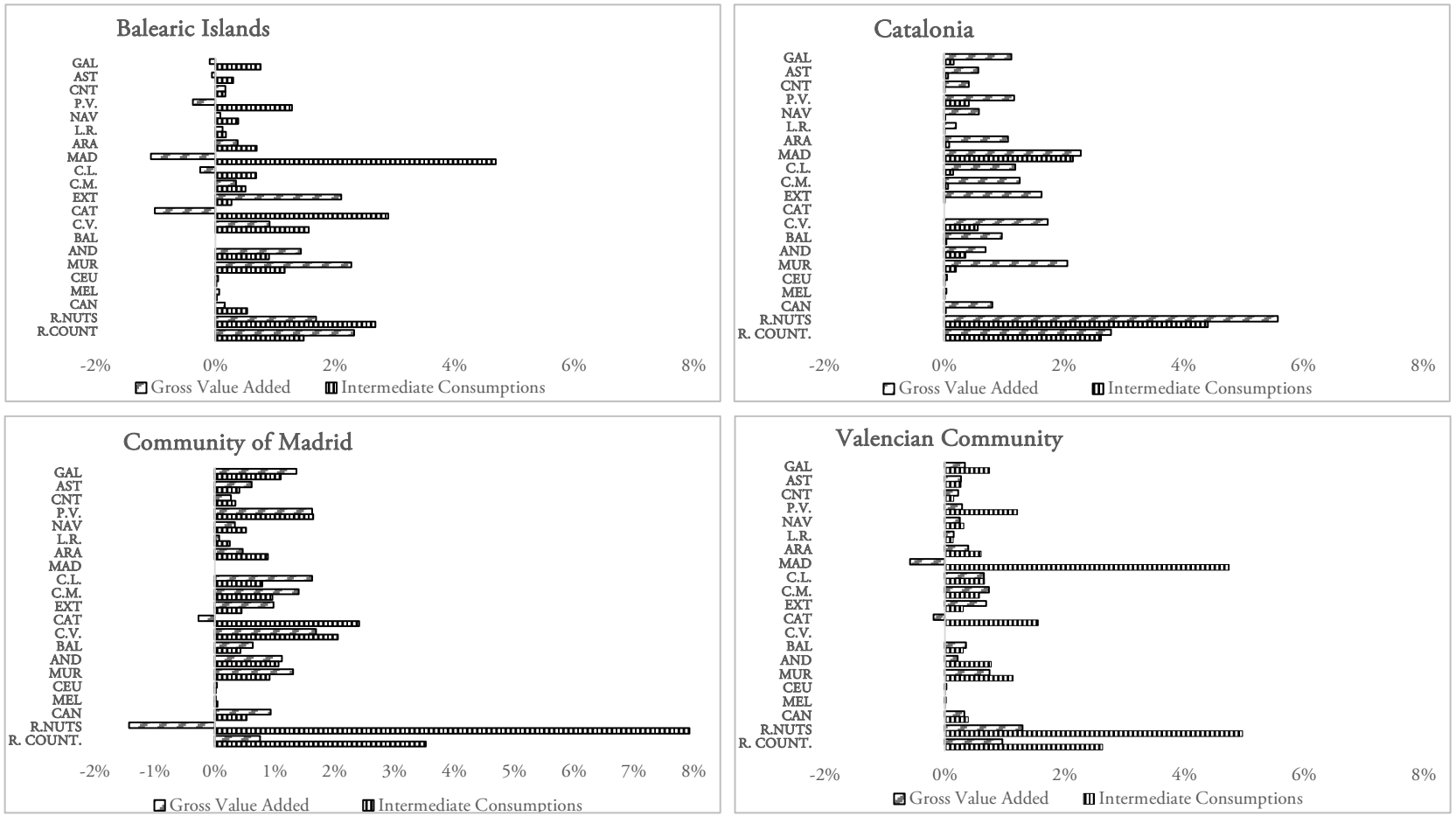
Finally, the interregional production structure of the four regions considered, see Figure 4, shows that the Balearic Islands present a negative GVA flow with the communities of Madrid, Catalonia, and the Basque Country. However, it should be noted that most of the company headquarters are in the Community of Madrid, Catalonia and the Basque Country. The Community of Madrid presents a considerably negative GVA flow with the rest of EU NUTS II, and a more moderate negative flow with Catalonia. In contrast, all Catalonia GVA flows are positive. Finally, in the Valencian Community, this region presents a negative flow with Catalonia and a much more intense negative flow with the Community of Madrid.

FIGURE 3.
Structure of intraregional intermediate consumptions and GVA of the Balearic Islands, Catalonia, Community of Madrid, and the Valencian Community



Note: MAD = Community of Madrid; CAT= Catalonia; C.V.= Valencian Community; BAL= Balearic Islands.
Source: compiled by the authors.

FIGURE 4.
Structure of interregional intermediate consumptions and GVA of the Balearic Islands, Catalonia, Community of Madrid, and the Valencian Community



Note: acronims are as in Figure 1.
Source: compiled by the authors.

4. Econometric gravity model

In order to justify and quantify the possible causes of interregional flows, an econometric gravity model is specified. Imbalances between actual flows (data) and potential flows under the hypothesis of independence (predicted or estimated flows) are linked causally through this gravity model. The proposed gravity model is inspired by the model put forward by Santamaría et al. (2023).

As discussed in Section 2, the gravity model is an adaptation of Newton's model. In this context, applying the gravity model yields a mathematical equation that relates interregional flows to the factors characterising the origin and destination regions of the flows.

In the case under study, the endogenous variable captures flow imbalances as represented by the elements of matrix **Y**. Each element y^{rs} of the square flow matrix **Y** is obtained from the following equation:

$$y^{rs} = \frac{x^{rs}}{x^{*rs}} \tag{3}$$

where y^{rs} is a generic element of matrix **Y**; x^{rs} is a generic element of the matrix of real interregional flows[7]; and x^{*rs} is a generic element of the matrix of interregional flows estimated under the independence hypothesis:

$$x^{*rs} = \sum_{r} x^{rs} \cdot \sum_{s} x^{rs} \tag{3.a}$$

The original data matrices are normalised, that is, they are weighted so that the sum of the elements of the row-transformed matrix is equal to one. Furthermore, in the event that any element of the flow matrix has a negative value, a rescaling of the elements of the original matrix is carried out so that, in the transformed matrix, all the elements of the interregional flow matrix are positive. From the rescaled matrix, the flow matrix estimated under the independence hypothesis is calculated.

Each of the elements of the matrix **Y**, that is, y_t^{rs}, measures, for year *t*, the deviation between the real flow from region *r* to region *s* and the estimated or predicted flow under the hypothesis of independence. A high value of the y_t^{rs} ratio would point to a greater imbalance in the flow between both regions.

In short, the gravity model considered is specified from the following function:

$$y_t^{rs} = \exp\left\{\sum_{m} \beta_m \, z_{mt}^{rs}\right\} \tag{4}$$

where z_{mt}^{rs} is a set of variables that determine the interregional flows between regions *r* and *s*, for year *t*, and β_m are the parameters that affect these *m* variables. z_{mt}^{rs} includes regressors representing the attraction between the origin and destination regions. Specifically, three variables capturing the intensity of interaction between regions have been used in this study: the geographical distance between regions, territorial contiguity (i.e., whether two regions share a common border) and an economic attraction variable. We define the economic attraction variable between regions *r* and *s* for year *t* as:

[7] For each year in the sample period (2000–2010), there are four matrices available: the matrix of intermediate consumption flows (Z), the matrix of final demand flows (F), the matrix of production flows (X), and the matrix of GVA flows (V). However, our analysis focuses on the GVA matrix, as it provides a more direct measure of the economic value generated in each region. Unlike intermediate or final demand flows, which may reflect reallocation or redistribution of value, GVA flows are better suited to capturing the net contribution of regional productive activities to the overall economy.

$$wg_t^{rs} = \frac{gdp_t^{\alpha r \cdot} \; gdp_t^{\beta . s}}{dist^{\gamma rs} \; \left(1 - \; \varphi \; contiguity^{rs}\right)} \tag{5}$$

where $\alpha = 1$, $\beta = 1$ and $\gamma = 1$. $gdp_t^{r \cdot}$ represents the GDP per capita of region *r*, the origin of the flow, for year *t*, while gdp_t^{s} represents the GDP per capita of region *s*, the destination of the flow, for the same year. $dist^{rs}$ measures the kilometer distance between regions *r* and *s*. When regions *r* and *s* share a border, $contiguity^{rs} = 1$. In this case, the distance between two contiguous regions is reduced by 10%, $\varphi = 0.1$ In short, the numerator of equation **(5)** captures the economic mass or attractiveness of the origin and destination regions, while the denominator quantifies the friction or barriers to interaction them.

According to Feenstra (2015) and Peeters & Chasco (2013), gravity model estimators can suffer from bias problems. To address such problems, it is recommended to include variables that capture factors affecting trade flows from both source and destination regions. Consequently, the proposed model incorporates dummy variables that capture the fixed effects of interregional flows.

From equation **(4)** the behavioural equation or econometric model **(6)** is specified. To do this, logarithms are taken in equation **(4)**, including the characteristics of the origin and destination regions of the interregional flows, and a random error term is added. Moreover, in order to analyse the effect of headquarters location on the intensity of flows, we incorporate a proxy variable H^r . The gravity model is specified through the following behavioural equation:

$$\log(y_t^{rs}) = \sum_r^R \gamma^{Or} \; DS^r \; + \sum_s^S \gamma^{Ds} \; DD^s \; + \delta \; T_t \; + \sum_m \beta_m \; z_{mt}^{rs} \; + \theta H^r \; + \; u_t^{rs} \tag{6}$$

where:

y_t^{rs} are the imbalance flows between region *r* and *s* at time *t*.

DS^r are dummy variables that represent the fixed effects of the region of origin of the flows.

DD^s are dummy variables that represent the fixed effects of the destination region of the flows.

T_t is a variable that includes possible temporal changes throughout the sample period.

z_{mt}^{rs} are variables that represent characteristics of the regions of origin of the flows, those of destination or simultaneously of both regions at time *t*.

H^r is a proxy variable the location.

u_t^{rs} is a random variable.

As control variables, a temporal variable and dummy variables representing the fixed effects of the origin and destination regions of interregional flows are included. As explanatory factors, the distance in kilometres between the origin and the destination regions[8], a dummy variable for the regional contiguity, and the economic attraction variable, see equation **(5),** are considered. The parameter θ captures the effect of the location of company headquarters on the intensity of interregional flows.

The variable amount of VAT is used as a proxy variable to represent the location effect of the company headquarters, see Appendix B. VAT is a consumption tax collected by the selling company at the point of sale. Although the tax is collected where the transaction with the final customer takes place, it is reported and paid to the tax authorities by the company's headquarters. As a result, from an administrative and statistical perspective, VAT revenues are typically recorded in the region where the company is fiscally registered, which usually corresponds to the location of its headquarters. Therefore, VAT records serve as a reasonable proxy for the geographical location of company headquarters.

To account for the dynamic nature of flows between regions, the gravity model specified in equation **(6)** is extended by including the lagged dependent variable, $\log(y_{t-1}^{rs})$. This inclusion captures potential persistence in interregional flows over time. The resulting behavioural equation is given by:

[8] The distance between the Spanish regions is obtained from the website distance.to.

$$\log(y_t^{rs}) = \sum_r^R \gamma^{Or} DS^r + \sum_s^S \gamma^{Ds} DD^s + \delta T_t + \sum_m \beta_m z_{mt}^{rs} + \theta H^r + \rho_1 \log(y_{t-1}^{rs}) + u_t^{rs} \quad (7)$$

where the parameter ρ_1 quantifies the *temporal inertia* of imbalances between actual and predicted trade flows under the independence hypothesis.

Finally, following the approach of Anselin (1988, 2003), the possible *spatial inertia* (i.e., substantive spatial autocorrelation) is introduced into the gravity model through the lagged endogenous variable, weighted by the economic attraction variable, $wg_{t-1}^{rs}.\log(y_t^{rs})$. In this way the following behaviour equation is obtained:

$$\log(y_t^{rs}) = \sum_r^R \gamma^{Or} DS^r + \sum_s^S \gamma^{Ds} DD^s + \delta T_t + \sum_m \beta_m z_{mt}^{rs} + \theta H^r + \rho_2 wg_{t-1}^{rs}.\log(y_{t-1}^{rs}) + u_t^{rs} \quad (8)$$

The parameter ρ_2 quantifies the *spatial inertia* of imbalances between actual and predicted trade flows under the independence hypothesis.

5. Application of the model to the Spanish regions

Based on the full gravity model specified in equation **(6)**, we analyse a set of alternative model specifications, presented in Table 3 and referred to as Model 1 to through 4.

We must note that the term $\log(y_t^{rs})$ into the gravity model proposed in equation **(6)**, is an endogenous variable and it is the logarithm of the imbalance of interregional flows between region *r* and region *s* over the sample period *t* and is denoted by. This variable was derived from the interregional GVA flows.

Analysing the results in Table 3, Models 1 to 4 show that the coefficient associated with the contiguity variable in the gravity model is statistically significant. Furthermore, the expected sign of this coefficient is negative, indicating that regions sharing a border tend to experience lower imbalances in interregional flows.

These models also show that the coefficient associated with the economic attraction variable[9] is statistically significant. The coefficient is also negative, suggesting that greater attraction between regions is associated with reduced interregional imbalance. Overall, it can be concluded that the economic attraction variable captures commercial flows between regions better than the contiguity variable, as the models including economic attraction exhibit greater accuracy.

Finally, when comparing the elasticities of these variables, the results are similar regardless of whether the contiguity or attraction variable is used. In fact, the elasticity in the first case (contiguity in Model 1) is *exp(-0.599)=0.549,* while in the second case (attraction in Model 2) it is *exp(-0.658)=0.518*. This suggests that the findings are robust.

When both contiguity and attraction are included as explanatory variables, the results confirm that the corresponding coefficients are statistically significant and have the expected sign, see Model 3 in Table 3. The combined elasticity is estimated at 31.6% (calculated as *exp(-0.634)·exp(-0.520)=0.316*).

[9] Given the high correlation among the distance between regions and the attraction variable, the latter variable has been used in this study.

Table 3.
The Gravity Model

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coefficient	t statistic	Coefficient	t statistic	Coefficient	t statistic	Coefficient	t statistic	Coefficient	t statistic	Coefficient	t statistic
Constant	-0.514***	-16.729	-0.512***	-16,922	-0.512***	-16.866	-0.208**	-1.991	-0.005	-0.095	-0.208*	-1.840
Contiguity	-0.599***	-5.877			-0.634***	-8.867	-0.622***	-8.307	-0.058	-1.551	-0.539***	-4.408
Economic attraction			-0.658***	-8,772	-0.520***	-4.761	-0.522***	-4.791	-0.065	-0.635		
Log (VAT)							-0.085***	-3.000	-0.009	-0.724	-0.075***	-2.509
Temporal inertia									0.897***	20.497		
Spatial inertia											0.185***	5.959
R squared	0.374		0.393		0.399		0.400		0.876		0.400	
F statistic	52.105		56.476		56.362		55.166		410.082		44.300	
Log. Likelihood	-3,219.688		-3,170.603		-3,154.435		-3,150.914					
J statistic									0.000		0.000	
Number of observations	3,179		3,179		3,179		3,179		2,601		2,601	

Note: The dependent variable is $\log(y_t^{rs})$. Temporal inertia=$\log(y_{t-1}^{rs})$; Spatial inertia=$wg_{t-1}^{rs} . \log(y_{t-1}^{rs})$All models include a time dummy variable and fixed effects of origin and destination regions. Models wihout *inertia* have been estimated by OLS and models with *inertia* have been estimated by Arellano and Bond (1991) method. The variance and covariance of the coefficients are adjusted for heteroscedasticity and autocorrelation. The estimations have been made using Eview Version 8.1 Software. The symbols *, ** and *** indicate the probabilistic significance levels of 10%, 5% and 1%, respectively.
Source: compiled by the authors.

Model 4 in Table 3 incorporates a variable representing the effect of company headquarters location. For this purpose, a proxy variable is included in the model, defined as the logarithm of the VAT collected in each region. The negative sign of the coefficient indicates that the *headquarters effect* reduces interregional imbalances between the regions involved. Thus, in regions where company headquarters are concentrated, the imbalances between the real and the estimated flow are smaller. Therefore, the presence of *headquarters effect* contributes to the overvaluation of regional GVA accounting, and consequently, of regional GDP.

Model 5 in Table 3 includes *temporal inertia* using equation **(7)**, and the goodness of fit improves substantially. However, the coefficient associated with the inertia term is close to one and highly significant, which may indicate a potential misspecification of the model. While the *temporal inertia* factor helps to explain the persistence of trade imbalances in GVA flows, its magnitude also suggests that some relevant variables might be omitted from the model.

To address potential endogeneity issues arising from the inclusion of the lagged dependent variable, the model has been estimated using the Arellano and Bond (1991) method. This approach applies the generalized method of moments (GMM), using lagged values of the endogenous variable as instruments. This estimation technique controls for the endogeneity of the lagged dependent variable and eliminates unobserved individual effects, improving the reliability of the estimates.

Lastly, Model 6 in Table 3 presents the results of the estimations of equation **(8),** which consider possible *spatial autocorrelation or spatial inertia*. To address potential endogeneity problems, the proposed gravity models are again estimated using the method proposed by Arellano and Bond (1991). The results indicate that that the explanatory variables related to *contiguity*, *spatial inertia* and the *headquarters effect* are all statistically significant, and their signs are consistent with expectations. Both the *contiguity* and the *headquarters effect* present negative values, while the coefficient for *spatial inertia* (spatial autocorrelation) is positive. However, it is worth noting that the inclusion of *spatial inertia* does not lead to an improvement in the model's goodness of fit compared to Model 4 in Table 3.

The results show that the *contiguity effect,* the *attraction effect,* and the *company headquarters effect* are all significant determinants of interregional imbalances in GVA flows. The positive *spatial inertia* coefficient supports the hypothesis that trade imbalances tend to persist across neighbouring regions, potentially due to structural or institutional interdependencies. Nonetheless, spatial autocorrelation may also reflect omitted spatially correlated factors, introducing a potential source of bias. In conclusion, the gravity model confirms the relevance of company headquarters location and spatial proximity —through contiguity and attraction— in explaining the GVA distribution across regions, and consequently, their GDP.

6. Conclusions

The analysis of interregional flows provided by the origin-destination matrices of the Input-Output Table is an appropriate tool for studying trade imbalances between regions, which ultimately determine economic inequalities.

One of the contributions of this study is the use of the accounting method to obtain the GVA flow matrix. This method provides a more accurate view of interregional flows because it allows for the correct allocation of GVA in cases where company headquarters and production centres are located in different regions. In such situations, the accounting method assigns the value added generated to the region where the economic activity actually takes place, rather than to the location of the administrative headquarters, thus avoiding distortions in the regional distribution of GVA.

A simple visualisation of the GVA flow matrix, which presents positive and negative values, provides an idea of the possible imbalances in trade flows between the different regions that make up the Spanish economy. It should be noted that the traditional techniques used in the analysis of interregional flows are insufficient to fully explain their behaviour, as they fail to account for imbalances that occur from flows being generated in one region but computed in another.

The fact that some of the interregional GVA flows are positive and others negative implies that some regions benefit from trade flows while others lose out. From an accounting point of view, this leads to an underestimation (or overvaluation) of the regional GVA and, consequently, of the regional GDP.

The estimated gravity model corroborates the hypothesis that the *contiguity effect* and *attraction effect* play a significant role in explaining the intensity of interregional trade flows. Specifically, better results are obtained when analysing the intensity of interregional GVA flows, leading to the conclusion that contiguity and stronger attraction between regions reduce trade imbalances between them.

The existence of a company headquarters effect is also confirmed. The results show that regions with a concentration of company headquarters experience smaller commercial imbalances. However, this effect also contributes to an overestimation of the regional GVA and, consequently, of the regional GDP, as value added generated in other regions may be statistically attributed to the region where the headquarters of the companies are located. While this may reflect the administrative structure of companies, from a territorial perspective, it can lead to distortions in regional economic indicators and reinforce existing disparities between regions.

To the extent that these distortions reinforce the position of already more developed regions and weaken the position of more peripheral or productive regions without a company headquarters, they may exacerbate economic imbalances between regions, thus hindering territorial convergence.

Finally, the results confirm the relevance of the location of company headquarters, as well as territorial *contiguity,* and economic attraction in determining the GVA flows for each region and, consequently, their GDP.

The analysis has been based on aggregate data, and the use of disaggregated sectoral data would allow for examining whether the effect of *company headquarters* differs depending on the type of economic activity. This could be an interesting direction for future research.

References

Anderson, J.E. (2010). *The gravity model*. NBER Working Paper No. 16576

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic.

Anselin, L. (2003). Spatial Econometrics. In Badi H. Baltagi (Ed.), *A Companion to Theoretical Econometrics* (pp. 310-330). Blackwell Publishing Ltd.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and application to employment equations, *Review of Economic Studies*, *58*(2), 277-297.

Baldwin, R., & Taglioni, D. (2006). *Gravity for dummies and dummies for gravity equations*. NBER Working Paper Series, Working paper 12516.

Distance.to. https://es.distance.to. Retrieved September 12, 2023.

EUROSTAT (1995). *Métodos de Contabilidad Regional. Valor Añadido Bruto y Formación de Capital Fijo por Rama de Actividad*. Brussels.

Evenett, S., & Keller, W. (2002). On Theories Explaining the Success of the Gravity Equation, *Journal of Political Economy, 110*(2), 281-316.

Feenstra, R.C. (2015). *Advanced international trade: theory and evidence*. Priceton University Press.

Frankel, J., & Romer, D. (1999). Does Trade Cause Growth? *American Economic Review*, *89*(3), 379-399.

Ghosh, A. (1958). Input–output approach in an allocation system, *Economica*, *25*(97), 58-64. https://doi.org/10.2307/2550694

He, L., & Duchin, F. (2009). Regional Development in China: Interregional Transportation Infrastructure and Regional Comparative Advantage, *Economic Systems Research*, *21*(1), 3-22.

Head, K., & Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. In *Handbook of International Economics*, *4*, 131-195. Elservier.

Head, K., & Mayer, T. (2019). Brands in Motion: How Frictions Shape Multinational Production, *American Economic Review, 109*(9), 3073-3124.

Head, K., & Mayer, T. (2021). The United States of Europe: A Gravity Model Evaluation of the Four Freedoms, *Journal of Economic Perspective, 35*(2), 23-48.

Hernández García, M.C. (1997). *El valor añadido como indicador económico de la responsabilidad social de la empresa: una aplicación empírica*. Universidad de la Laguna.

Kabir, M., Salim, R., & Al-Mawali, N. (2017). The gravity model and trade flows: Recent developments in econometric modelling and empirical evidence, *Economic Analysis and Policy*, *56*, 60-71.

Leontief, W. (1953). *The structure of American economy 1919-1929: An empirical application of equilibrium analysis*. Harvard University Press.

Mayer, T., Vicard, V., & Zignago, S. (2019). The cost of non-Europe, revisited. *Economic Policy*, *34*(98), 145-199.

Nitsch, V., & Wolf, N. (2013). Tear down this wall: on the persistence of borders in trade, *Canadian Journal of Economics*, *46*(1), 154-179.

Official Journal of The European Union (2013). *Regulation (EU) No 549/2013 of the European Parliament and of the Council of 21 May 2013 on the European system of national and regional accounts in the European Union Text with EEA relevance*. Brussels.

Pavía Miralles, J.M., & Cabrer-Borrás, B. (2004). Espacio y comercio interregional en España. *Estudios de Economía Aplicada*. *22*(3), 517-538.

Peeters, L., & Chasco, C. (2016). Identifying local determinants of destination choices of international immigrants to the Madrid metropolitan area. *Papers in Regional Science*, *95*(2), 281-307. https://doi.org/10.1111/pirs.12130

Reuveny, R., & Li, Q. (2003). Economic openness, democracy, and income inequality. An Empirical Analysis, *Comparative Political Studies*, *36*(5), 575-601.

Santamaría, M., Ventura, J., & Yeşilbayraktar, U. (2023). Exploring European regional trade, *Journal of International Economics*, *146*. https://doi.org/10.1016/j.jinteco.2023.103747

Serrano-Domingo, G., Cabrer-Borrás, B., & Requena-Silvente, F. (2020). Networks and location of foreign migrants: evidence for Southern Europe. *Journal of Regional Research, 47*(2), 97-112. https://doi.org/10.38191/iirr-jorr.20.012

Shahriar, S., Qian, L., Kea, S., & Abdullahi, Z.M. (2019). The Gravity Model of Trade: A Theoretical Perspective, *Review of Innovation and Competitiveness, A Journal of Economic and Social Research*, *5*(1), 21-42.

Tinbergen, J. (1962). *Shaping the World Economy*. The Twentieth Century Fund.

Thissen, M., Lankhuizen, M., Van Oort, F., Los, B., & Diodato, D. (2018). EUREGIO: The construction of a regional IO DTABASE with regional detail for Europe for 2000-2010. Tinbergen Institute Discussion Paper, TI 2018-084/VI.

Thissen, M., Ivanova, O., Mandras, G., & Husby, T. (2019). *European NUTS 2 regions: construction of interregional trade-linked Supply and Use tables with consistent transport flows*. European Commission, JRC Technical Report.

Tormo, L., Rico, P., Morillas-Jurado, F., & Cabrer-Borrás, B. (2023). A new approach to the hypothetical extraction method: regional full extraction. *Economic Systems Research*. https://doi.org/10.1080/09535314.2022.2157249

Winters, A., McCulloch, N., & McKay, A. (2004). Trade Liberation and Poverty: The Evidence so Far, *Journal of Economic Literature*, *42*(1), 72-115.

Wang, Z. (2021). Headquarters gravity: How multinationals shape international trade. *Journal of International Economics*, *131*. https://doi.org/10.1016/j.jinteco.2021.103477

Wrona, J. (2018). Border effects without borders: What divides Japan's internal trade? *International Economic Review*, *59*(3), 1209-1262.

Yotov, Y.V. (2021). On the role of domestic trade flows for estimating the gravity model of trade. *Contemporary Economic Policy*, *40*, 526-540.

ORCID

Francisco Morillas-Jurado https://orcid.org/0000-0002-4204-906X

Paz Rico-Belda https://orcid.org/ 0000-0002-5634-6022

Appendix A

GVA can be defined in the following two ways:

1. Remuneration of production factors; employee remuneration, consumption of fixed capital, interest, and business profits. Thus, the GVA, in a simplified way, is obtained as the sum of the wage bill plus the gross operating surplus.

In the case of a regional input-output matrix, there may be an accounting problem when the headquarters of company A is in region *r*, which does not coincide with region *s* where its production plant is located (see EUROSTAT, 1995).

2. Value of the set of goods and services produced in a region during a period of time, discounting indirect taxes and intermediate consumption.

   From the accounting point of view, the value of GVA can be negative when production is less than its intermediate consumption (IC). In the case of a regional input-output matrix, an accounting problem may exist when region *s* where production is located does not coincide with region *r* from which the IC comes (see Hernández García, 1997).

There are different ways to calculate GVA flows. Below are just some examples:

The **Expenditure Method** consists of measuring the total expenses (total purchases) that are made in a region in consumption and investment by economic agents (families, companies and government). Purchases of intermediate goods and/or services are excluded.

The **Income Method** consists in measuring the remuneration of the factors of production (workers, entrepreneurs, capital, shareholders, etc.).

The **Value Added Method** consists in subtracting intermediate consumption from the total value of a region's production.

GVA = Production – IC

Example of accounting method for Regional GVA

a) **Production approach.** See the Official Journal of the European Union (2013). Regulation No. 549/2013; section 13.35.

Suppose we have a company that has two centres, the headquarters (local EAU), in Region 1 and the production plant (local EAU) in Region 2. This Company A has a production volume of 90 €, which the Central Headquarters (local EAU) manages, located in Region 1. In this case the income is accounted for as:

Production (Sales) of Company A			
	Region 1	**Region 2**	**TOTAL**
Region 1	90 €		90 €
Region 2			
TOTAL	90 €		90 €

The value of the company's intermediate consumption (IC), in total, amounts to 24 €, while in the production plant located in Region 2 this is 20 €, and the IC at the headquarters located in Region 1 amounts to 4 €. The expenses of Company A in IC are accounted for as:

IC of Company A			
	Region 1	**Region 2**	**TOTAL**
Region 1	4 €		4 €
Region 2		20 €	20 €
TOTAL	4 €	20 €	24 €

Furthermore, the amount of personnel costs (PC) at the production plant located in Region 2 amount to 35 €, while the PC costs at the headquarters located in Region 1 amount to 6 €. These expenses in PC are accounted for as:

PC costs of Company A (GVA component)			
	Region 1	**Region 2**	**TOTAL**
Region 1	6 €		6 €
Region 2		35 €	35 €
TOTAL	6 €	35 €	41 €

The gross operating surplus (remuneration of capital) at the production plant located in Region 2 amounts to (-55 €), while at the headquarters located in Region 1 this amounts to 80 €. These concepts are accounted for as:

Gross Operating Surplus of Company A (GVA component)			
	Region 1	**Region 2**	**TOTAL**
Region 1	80 €		80 €
Region 2		-55 €	-55 €
TOTAL	80 €	-55	25 €

To obtain the GVA of Company A in a consolidated manner, through the difference between its production and its IC, we have:

GVA of Company A			
	Region 1	**Region 2**	**TOTAL**
Region 1	86 €	-20 €	66 €
Region 2			
TOTAL	86 €	-20 €	66 €

The previous tables show that the total GVA of Company A is €66 and that of Region 1 is 86 €. However, there is a negative GVA flow, from an accounting point of view, between Region 1 and Region 2 that amounts to (-20 €).

b) **Income approach.** See the Official Journal of the European Union (2013). Regulation No. 549/2013; section 13.36.

An alternative way of calculating the GVA (following the methodological recommendations of regional accounting; see EUROSTAT, 1995) assumes that Company A has a production volume of 90 €, which is managed by the headquarters of Company A in Region 1. The PC of the production plant, located in Region 2 amounts to 35 € while the headquarters, located in Region 1, registers a PC of 6 €. The total costs in PC of Company A amount to 41 €. The PC of the production plant located in Region 2 is paid by the headquarters in the amount of 35 €. The company's total operating surplus (profit) amounts to 25 €. Thus, the company's total GVA amounts to 66 € (41 €+25 €). The total GVA to be accounted for by Company A's headquarters is: 86 € (66 €+20 €), while the GVA flow between Region 1 (headquarters) and Region 2 (production plant) is (-20 €).

GVA of Company A			
	Region 1	**Region 2**	**TOTAL**
Region 1	66 €	-20 €	46 €
Region 2	20 €		20 €
TOTAL	86 €	-20 €	66 €

Appendix B: Sources of data on VAT

Several data sources were used to obtain the VAT accounting of the Spanish regions:

- **Community of Navarre, special regime:**
  Gráfico nº 11. Evolución temporal de la recaudación tributaria - memoria 2019 - navarra.es
- **Spain, general regime:**
  https://javiersevillano.es/IVA.htm#Recaudaci%C3%B3n%20Del%Microsoft Word - nota_calculo_indicador.docx (ine.es)
- **Canary Islands, special regime:**
  ISTAC | Recaudación Tributaria del Estado acumulada. Comunidades autónomas y provincias por periodos | Banco de datos (gobiernodecanarias.org)

  https://www3.gobiernodecanarias.org/istac/statistical-visualizer/visualizer/data.html?resourceType=data
- **Basque Country, special regime:**
  Recaudación de las Diputaciones Forales por tributos concertados - Departamento de Hacienda y Economía - Gobierno Vasco - Euskadi.eus

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 91-112
https://doi.org/10.38191/iirr-jorr.25.003

Articles

Impacto económico regional de las Universidades latinoamericanas. Un análisis comparado

Jordi Suriñach[*], *Esther Vayá*[**], *Jordi López-Tamayo*[***], *Joaquim Murillo*[****]

Recibido: 31 de julio de 2024
Aceptado: 25 de marzo de 2025

Resumen:

El artículo supone la primera aportación en la que se estima el impacto económico de un conjunto de nueve universidades de América Latina siguiendo un enfoque de demanda y con una metodología homogénea. Las universidades analizadas responden a diferentes perfiles en función de su titularidad, dimensión, antigüedad o ubicación en el territorio. Aplicando la metodología input-output se estiman los impactos derivados de los gastos e inversiones que realizan tanto las universidades como los agentes vinculados a ellas. Los resultados muestran unos impactos significativos (del 4% del PIB regional dependiendo del caso), que se extienden por todo el territorio nacional y todos los sectores económicos. Se demuestra que el impacto económico y social es muy relevante, sobre todo en las universidades con un alto nivel de actividades de I+D+I.

Palabras clave: Impacto económico universidades; desarrollo regional y local; metodología Input-Output; América Latina.
Clasificación JEL: E24; I23; L89; R15; R58.

Regional economic impact of Latin American Universities. A comparative analysis

Abstract:

The article represents the first contribution in which the economic impact of a group of nine Latin American universities is estimated, following a demand-based approach and a homogeneous methodology. The analyzed universities have different profiles depending on their ownership, size, age, or geographic location. Using the input-output methodology, the impacts derived from the expenditures and investments made by both the universities and their related agents are estimated. The results show significant impacts (up to 4% of the regional GDP, depending on the case), which extend across the entire national territory and all economic sectors. The study demonstrates that the economic and social impact is highly relevant, providing a substantial return, overall in universities with a high level of R+D+I activities.

Keywords: University economic impact; regional and local development; Input-output methodology; Latin America.
JEL Classification: E24; I23; L89; R15; R58.

[*] Grupo de Investigación AQR, Universitat de Barcelona. España. jsurinach@ub.edu
[**] Grupo de Investigación AQR, Universitat de Barcelona. España. evaya@ub.edu
[***] Grupo de Investigación AQR, Universitat de Barcelona. España. jlt_lopez@ub.edu
[****] Grupo de Investigación AQR, Universitat de Barcelona. España. j.murillo@ub.edu
Autor para correspondencia: jsurinach@ub.edu

1. Introducción

Según datos de 2021 de CEPAL (CEPALSTAT), el conjunto de países de América Latina y el Caribe presentaron un producto interior bruto per cápita (PIB pc) promedio de 18.560 $, cifra inferior a la media mundial de 20.271 $ y bastante alejada de los 38.168$ de media de países de la OCDE. A ese menor nivel de desarrollo, según se recoge en el informe anual de la CEPAL (2024), se le une una trampa de baja capacidad para crecer, para lo cual es preciso desarrollar políticas centradas en incrementar la capacidad productiva por medio de un aumento de la productividad, del fomento de la inversión y de la generación de empleo de calidad. Adicionalmente, es una región con un idioma mayoritario (castellano) y lazos culturales elevados.

En este sentido, las instituciones universitarias impactan en el crecimiento económico a medio/largo plazo en las regiones/países en las que están ubicadas. Ello se debe a que, haciendo uso de un concepto económico como es la función de producción de una economía, la triple dimensión de las actividades que realizan las universidades (docencia, investigación y transferencia de conocimiento), llevan a que se incremente el capital tecnológico (progreso técnico), el empleo y la dotación de capital humano presentes en la región/país, favoreciendo su crecimiento.

Pero de forma similar, las universidades también impactan en el crecimiento económico más a corto plazo en la medida en la que son claros agentes generadores de gasto e inversión en los territorios en los que se encuentran. A su vez, existen otros agentes ligados a las universidades y que son también generadores de gasto y, por tanto, estimuladores del aumento de actividad de un territorio, entre los que destacarían los estudiantes de las mismas, todo el profesorado e investigadores, las empresas *starts-up* y *spin-off* o los parques científicos vinculados con las universidades, etc.

Teniendo ello en cuenta, las universidades se revelan por tanto como unos agentes claves para favorecer el crecimiento económico del que precisa América Latina (Moreno-Brid y Ruiz-Nápoles, 2009; Hernández et al, 2015). Si bien son numerosos los trabajos en los que se analiza el impacto de las universidades para diversas geografías, no existen, hasta donde conocemos, aportaciones que realicen un análisis homogéneo, utilizando la misma metodología y para un mismo año, del impacto de las universidades latinoamericanas.

Por todo ello, y con el propósito de cubrir dicho déficit, el presente artículo tiene como objetivo analizar el impacto económico en el corto plazo de un conjunto de universidades de América Latina sobre su territorio. La aportación del artículo es doble: primero, diseñar y aplicar una metodología de cuantificación que permita medir el impacto económico universitario a corto plazo; y, segundo, obtener, por primera vez, una estimación comparada del impacto universitario para una muestra destacada de universidades latinoamericanas que pueda constituir el punto de partida para posteriores estudios comparados del impacto universitario en América Latina.

La estructura del artículo es la siguiente. Tras la introducción, en la sección 2 se presenta una revisión de la literatura de medición del impacto universitario. En la sección 3 se describe la metodología utilizada. En la sección 4 se presentan las características de las universidades analizadas mientras que en la sección 5 se presentan los principales resultados. Por último, la sección 6 sintetiza las principales conclusiones del estudio.

2. Aproximaciones al estudio de impacto económico de las Universidades

En los últimos años, se han realizado diversas contribuciones a la estimación del impacto de las universidades, tanto en el largo como en el corto plazo. Así, siguiendo lo que se denomina un "enfoque de oferta", han sido diversos los autores que han estudiado el impacto a medio y largo plazo que las universidades tienen sobre el crecimiento económico (Blackwell et al., 2002; Drucker y Goldstein, 2007; Pastor et al, 2016). Algunos de ellos analizan el impacto de las universidades por medio de un análisis contrafactual donde se estudia qué habría ocurrido en términos de capital humano generado, capital

tecnológico, tasas de paro o de actividad en caso de no haber existido una universidad concreta (Pérez et al, 2015) En otros casos, se recurre a la modelización econométrica de un modelo de crecimiento basado en una función de producción ampliada que permita aislar el efecto de las universidades (Segarra et al, 2011; Valero y Van Reenen, 2019). De manera general, en dichos trabajos se precisa de información de variables de stock de capital físico, de capital humano y de capital tecnológico para períodos temporales extensos y/o múltiples entornos geográficos. La no disponibilidad de esta información para la mayoría de economías latinoamericanas supone una limitación importante para la estimación del impacto económico a largo plazo de las universidades siguiendo este enfoque de oferta.

Otra parte de la literatura ha cuantificado el impacto que producen las universidades en la economía en el corto plazo ("enfoque de demanda") medido en términos de puestos de trabajo, producción y renta generada por todos los gastos e inversiones realizados por dichas universidades en el ejercicio de su actividad. En este enfoque, el impacto económico se cuantifica habitualmente a partir de tres componentes: impacto directo (consecuencia directa de los gastos e inversiones realizadas por la universidad), indirecto (derivado de las relaciones intersectoriales que se generan fruto de los flujos de bienes y servicios que requieren las empresas generadoras de impacto directo para poder desarrollar su actividad), e inducido (el aumento de la producción derivada de la parte de las rentas del trabajo, generadas directa e indirectamente, que se destinan al consumo de bienes y servicios).

Los primeros estudios de impacto económico de las universidades vía demanda surgieron en Estados Unidos en la década de los sesenta del siglo pasado y se fueron consolidando, metodológicamente, en la década siguiente a partir de la aportación de Caffrey e Isaacs (1971). En la aportación de estos autores es donde se origina el modelo ACE (American Council on Education) y a partir de aquí han ido apareciendo derivaciones del mismo. El modelo ACE fue el método de cuantificación de referencia del impacto de las universidades norteamericanas hasta principios de los noventa.

No obstante, como alternativa a este tipo de métodos de estimación, en que los impactos indirectos se calculan con multiplicadores, se propusieron otros métodos de estimación del impacto indirecto entre los que destacan los basados en la utilización de las tablas input-output (TIO) de las economías que permiten aflorar así la actividad derivada de las demandas de inputs y consumos intermedios de la universidad en otros sectores (Miller y Blair, 2009). Una de las principales ventajas de aplicar la metodología input-output, en comparación al modelo ACE, es que requiere de menos información primaria y menos hipótesis para estimar la actividad derivada de las relaciones intersectoriales con proveedores. Así, a partir de las TIO de cada país, la metodología input-output permite aproximar la totalidad de impactos indirectos (y también inducidos).

La metodología input-output se ha generalizado en este tipo de estudios de impacto, durante los últimos años, en los que se ha tratado de estimar el impacto económico en el corto plazo de universidades concretas (D'Allegro y Paff, 2010; Garrido-Yserte y Gallo-Rivera, 2010; Kelly et al., 2014; Tripp Umbach, 2018; Nauffal, 2019; o Suriñach et al., 2017, 2021, 2023). No obstante, como se mencionó previamente, no se ha encontrado en la literatura un enfoque uniforme que aplique la misma metodología para un año determinado para evaluar el impacto de diferentes universidades en América Latina, hecho que fundamenta el objetivo del presente artículo.

3. Detalle metodológico aplicado para la estimación del impacto económico de las universidades latinoamericanas

En el presente artículo se procede a hacer una primera aportación en la cuantificación del impacto económico en el corto plazo de las universidades latinoamericanas siguiendo un enfoque de demanda. Para ello se han seleccionado un total de nueve universidades: Universidad CES y Universidad de la Costa (CUC), en Colombia; Universidad de Chihuahua (UACH) y Universidad Autónoma de Chiapas (UNACH), en México; Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA), en Argentina; Universidad de Chile (UCH) y Universidad de Tarapacá (UTA), en Chile; y Universidad Nacional de Costa Rica (UCR) y Universidad Tecnológica de Panamá (UTP), de la región centroamericana.

La selección final de las universidades analizadas responde a diversos factores. Así, se priorizó que fueran universidades cuyos países dispusieran de TIO homogéneas para poder estimar los impactos indirectos e inducidos correspondientes. A su vez, se consideró importante disponer de una muestra de universidades que captaran una variedad de características: la titularidad (públicas como privadas); la ubicación (en la ciudad capital del país y en regiones cercanas a áreas fronterizas); la dimensión (tanto en términos del tamaño de la población total de la región y de las tasas de población indígena, como del número de estudiantes); la antigüedad; la ubicación territorial de las mismas (en base a la presencia física de sus sedes y unidades regionales); el número de programas académicos; o la inclusión o no en el Times Higher Education Impact Ranking (una de las clasificaciones mejor valoradas a nivel internacional). Si bien este conjunto de universidades no constituye una muestra representativa de la totalidad de universidades de Latinoamérica (imposible dado el elevado número de universidades existentes y los requisitos de información que son necesarios para la estimación del impacto de cada una de ellas), sí que se ha buscado captar al máximo la diversidad de perfiles de universidades presentes[1].

La cuantificación del impacto económico de dichas universidades se ha realizado para el año 2019 (último año antes de la irrupción de la pandemia del SARS-CoV-2) y en términos de las siguientes macromagnitudes: Facturación, PIB, Población Ocupada, Remuneraciones salariales y rendimientos fiscales.

El impacto total estimado es el resultado de la agregación de tres tipos de efectos: directo, indirecto e inducido. Para la cuantificación del impacto total se siguieron las siguientes fases.

I. Determinación de los agentes generadores de impacto.

En primer lugar, se ha requerido aflorar el conjunto de agentes generadores de impacto y que están vinculados al entorno universitario. El primer agente es, obviamente, la propia universidad. Sin embargo, éste no es el único agente que realiza un gasto o actividad que tiene efectos económicos claros. También deben considerarse a los estudiantes universitarios en función de su lugar de residencia y a todos aquellos visitantes de las universidades que generan un gasto en la región/país donde está ubicada la universidad (amigos y/o familiares que visitan a los estudiantes y al Personal Docente e Investigador-PDI), así como también los asistentes a los eventos organizados por ellas como congresos, jornadas, ...). A su vez, además de los tres agentes anteriores, también debe considerarse un cuarto: el conjunto de entidades cuya actividad se encuentra vinculada a la universidad, como Parques Científicos y Tecnológicos u otras instituciones/empresas como pueden ser fundaciones o consorcios (todos ellos entes vinculados a la universidad, pero con identificador fiscal propio). Por último, el quinto agente considerado lo ha formado el conjunto de empresas cuya actividad tiene su origen en la existencia de la propia universidad (como las empresas ubicadas en los Parques Científicos y Tecnológicos, las *start-ups* y *spin-offs* originadas), las empresas que son adjudicatarias de concesiones administrativas y las empresas que acogen estudiantes en prácticas y/o voluntariado de estudiantes y staff.

II. Recopilación de información para la estimación del impacto directo.

La recogida de información para el cómputo del efecto directo se realizó a través del envío a cada Universidad (tras una reunión previa con el equipo rectoral de cada universidad y múltiples reuniones con los servicios de gerencia y responsables de información de las universidades) de un cuestionario muy detallado en el que se les solicitaba información relativa a cinco bloques: Actividad económica que llevan a cabo todos los colectivos que operan con el identificador fiscal de cada universidad; Actividad económica de Instituciones/Empresas vinculadas a la universidad; Actividad empresarial imputable a la existencia de la universidad; Actividad económica generada por los visitantes que se desplazan a la ciudad donde se ubica la universidad, por motivos relacionados con ella (amigos y familiares que visitan a estudiantes y PDI, asistentes a congresos, jornadas, ...); y Actividad económica generada por los estudiantes, en función de su lugar de residencia.

[1] La selección inicial de universidades se realizó con el asesoramiento de expertos en el sistema universitario del Banco Interamericano de Desarrollo. En este sentido, es preciso decir también que al inicio del estudio se trató de incluir a otras universidades brasileñas y argentinas, pero desistieron de participar en el proyecto.

En todo este proceso, para facilitar el análisis comparativo, se requirió unificar las unidades monetarias con las que cuantificar las variables (dólares estadounidenses-USD) aplicando el tipo de cambio medio del año 2019.

Con dicha información se consiguió estimar el impacto directo generado por los distintos agentes, en términos de facturación/gasto, puestos de trabajo, rentas salariales y valor añadido bruto (variable que mide la contribución al PIB del país/región de cada universidad).

III. Cuantificación del impacto indirecto

Para la realización del análisis de impacto indirecto, se recopiló la información sobre las TIO disponibles en la OCDE para una desagregación sectorial a 36 sectores de actividad y para la práctica totalidad de países del estudio: Argentina, Chile, Colombia, Costa Rica y México (última edición de 2018 de las TIO a fecha de 2015[2]). Únicamente no se dispuso de la TIO de la OCDE para Panamá, utilizándose para dicho país la de Costa Rica (país con el que guarda mayores similitudes).

IV. Cuantificación del impacto inducido

Para estimar el impacto inducido se empleó la retribución media bruta asociada a los puestos de trabajo directos e indirectos estimados calculada a partir de datos de estadísticas oficiales detrayendo las cuotas medias estimadas de ahorro, impuestos y otras cuotas (también aplicando las tasas medias oficiales). La cantidad que retorna al sistema económico se repartió entre los distintos sectores a partir de la distribución del consumo de los hogares de la TIO.

V. Relativización territorial del impacto

Una vez estimados los tres impactos, se obtuvo el impacto total como agregación de los impactos directo, indirecto e inducido.

Tras ello, una dimensión importante del análisis consistió en relativizar el impacto en el territorio en el que se ubica cada sede o sedes de cada universidad. Para ello, se seleccionaron unas divisiones políticas-administrativas lo más homogéneas posibles entre países: Provincias en el caso de Argentina (24), Regiones en el caso de Chile (16), Departamentos en Colombia (34), Provincias en Costa Rica (7), Entidades Federativas para México (32) y Provincias para Panamá (10).

Una vez escogida la división territorial, se procedió a analizar, para cada universidad, en cuántos territorios tenían localizadas sus sedes. Para aquellas universidades que únicamente tenían presencia en uno de los territorios de su división (provincia, región, estado federativo o departamento), se realizó la siguiente asignación territorial de los impactos. Todo el impacto directo se asignó al territorio donde se ubicaba la sede. En relación con el impacto indirecto, se asignó el 65% del impacto indirecto de primer orden de los sectores primarios y secundarios y el 80% en el caso de los sectores de construcción y servicios al territorio donde se ubicaba la universidad. El resto del impacto indirecto se asignó sectorialmente al territorio donde estaba la sede en una proporción igual al peso que dicho territorio tenía en cada sector en términos de ocupación (o Valor Añadido Bruto, VAB) respecto el país[3]. Un procedimiento similar se aplicó para el caso de la distribución territorial del impacto inducido.

En cambio, para aquellas universidades que tenían presencia en más de un territorio, se siguió un procedimiento similar al anterior con algunas adaptaciones. Así, los impactos directo e indirecto de primer orden se distribuyeron entre los diferentes territorios donde estaban ubicadas las sedes de la universidad en función de la distribución por sedes del número de estudiantes. El impacto inducido de primer orden se distribuyó entre los territorios en los que había presencia de la universidad en función del peso que dichos territorios tenían, respecto el total nacional, en términos de la suma de las rentas directas e indirectas

[2] En el mes de noviembre de 2021, la OCDE publicó la última edición de 2021 de las TIO referidas al año 2018. A pesar de su mayor proximidad temporal, se desestimó su utilización dado que, a diferencia de las tablas de la edición anterior, no suministran información detallada (ni sectorialmente ni de manera agregada) del Valor Añadido Bruto ni de su desagregación por componentes. En https://stats.oecd.org/Index.aspx?DataSetCode=IOTS se proporcionan tablas para Argentina, Brasil, Chile, Colombia, Costa Rica y México, y otros países de la OCDE.

[3] Dichos porcentajes se aplicaron a partir de diferentes consultas realizadas con las universidades analizadas en relación al origen geográfico de sus proveedores.

generadas por la universidad. Para la distribución del resto de impactos se siguió el mismo procedimiento que el explicado anteriormente para el caso de las universidades con una sola sede.

Cabe notar que el utilizar una misma metodología para todas las universidades permite un análisis comparativo entre resultados asociados a todas ellas (y de distintos países). Para facilitar la comparación de las magnitudes monetarias y de personas entre universidades, se relativizaron las macromagnitudes estimadas respecto al número de estudiantes y al número de ocupados directos en la universidad.

4. Características de las universidades analizadas

Antes de proceder a presentar los resultados asociados al impacto económico generado por cada una de las nueve universidades latinoamericanas analizadas, en este apartado se presentan algunas de sus características básicas (cuadro 1).

Todas las universidades son públicas excepto las dos de Colombia (CES y CUC). Asimismo, algunas se ubican en una única región, provincia, departamento o estado federativo (UACH, UNACH, UNNOBA, UCH, CES) y otras están presentes en más de una (UTA, UTP, CUC, UCR).

Por otro lado, su dimensión es muy variable: desde universidades con un elevado número de estudiantes y ocupados (UACH, UCH o UCR), con más de 29.000 estudiantes y más de 4.000 ocupados, a otras de menor tamaño (CES, UNNOBA, UTA y CUC) que no llegan a los 14.000 estudiantes y a los 1.500 ocupados. Otras presentan un tamaño medio en estudiantes y/o en ocupados (UNACH, UTP).

Asimismo, las cuotas de mercado que representan las universidades, en el conjunto del sistema universitario de cada país es variable, dependiendo del tamaño de la universidad, pero también del país y de su estructura universitaria. En un extremo se encuentran UCR y UTP (con cuotas del 36% y 11,65%, respectivamente), y en otro las universidades mexicanas (UACH y UNACH) o argentinas (UNNOBA), con cuotas de mercado menores al 1%.

CUADRO 1.
Análisis comparativo entre universidades. Principales resultados

	Colombia	Colombia	Mexico	Chile	Costa Rica	Mexico	Argentina	Chile	Panama
	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
DATOS DEL PAIS									
PIB del país 2019 PRECIOS 2019 (MUSD)	323.616	323.616	1.268.871	282.318	61.801	1.268.871	451.815	282.318	66.801
Población 2019 (Millones de personas)	50,3	50,3	127,6	19	5	127,6	44,9	19	4,2
Ocupados 2019 (Millones de personas)	24,2	24,2	55,2	8,9	2,2	55,2	18,7	8,9	2
DATOS DE LA UNIVERSIDAD									
Publica/Privada	Privada	Privada	Pública	Pública	Pública	Pública	Pública	Pública	Pública
Sedes	Antioquia	Atlántico y Nariño	Chihuahua	Región Metropolitana de Santiago	Alajuela, Cartago, Guanacaste, Limón, Puntarenas, San José	Chiapas	Buenos Aires	Arica y Tarapacá	Todas las provincias menos Darien y Herrera
Estudiantes 2019	6.205	13.081	29.180	37.264	47.102	23.227	6.902	9.196	18.912
Ocupados ETC de la Universidad 2019	1.452	865	4.298	11.951	7.601	2.177	582	875	2.916
Estudiantes 2019/Ocupados ETC de la Universidad 2019	4,3	15,1	6,8	3,1	6,2	10,7	11,9	10,5	6,5
Titulados (solo de Grado) 2019	542	1.658	4.223	7.358	5.274	3.573	204	708	3.024

CUADRO 1. CONT.
Análisis comparativo entre universidades. Principales resultados

	Colombia	Colombia	Mexico	Chile	Costa Rica	Mexico	Argentina	Chile	Panama
	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
¿Existencia de Parque Científico?	No	No	Sí (Tecnoparque)	Sí	No	No	No	No	No
¿Empresas vinculadas con identificador fiscal propio?	Sí	No	Sí	Sí	No	Sí	Sí	Sí	Sí
¿Tiene *Start-ups, spin-offs*, etc?	Sí	No	No	Sí	Sí	Sí	No	No	Sí
Número de estudiantes que hacen prácticas o son voluntarios en términos de ETC	404	1.008	2.191	511	586	1.156	19	34	147
Cuota de la universidad respecto total graduados del país	1,00%	0,55%	1,00%	6,50%	36,00%	0,62%	0,20%	1,00%	11,65%

Fuente: Elaboración propia. **Nota**:(MUSD): Millones de dólares de Estados Unidos.

Adicionalmente, la ratio estudiante/ocupado ETC (Equivalente Tiempo Completo) es variable, situándose alrededor de 15 en CUC, de 10 en UNACH, UNNOBA y UTA, de 7 en UACH, UTP, UCR, o alrededor de 3,5 en UCH y CES. No se detecta una explicación asociada a estas diferencias, ni por tamaño, ni por titularidad (pública o privada), ni por regiones[4] en las que se ubica la universidad. Una posible explicación son las actividades adicionales a las estrictamente docentes que se desarrollan en la universidad (por ejemplo, en CES existe una elevada transferencia de conocimiento que puede justificar parte del "extra" de ocupados que serían los normales, o en varias universidades con bajo ratio estudiante/ocupado ETC hay ubicados parques científicos; asimismo, también en UCH se dispone de un parque científico).

Por último, las actividades adicionales a las estrictamente docentes también varían y justifican parte de los distintos resultados de impacto económico generado: algunas de ellas tienen una elevada actividad de transferencia de conocimiento y de servicios prestados conjuntamente, o no, con otras empresas (por ejemplo, CES y UTP, respectivamente), otras tienen un parque científico o tecnológico (por ejemplo, UACH y UCH), y la mitad de ellas (UNACH, UCH, UTP, CES y UCR) disponen de empresas nuevas de base tecnológica (*Start-ups* y/o *Spin-off*). UCH destaca por su actividad en prestaciones médicas y de otros servicios y la gran cantidad de Centros de investigación, científicos y tecnológicos, y organismos dependientes de que dispone, entre otros.

5. Resultados

Los cuadros 2 y 3 y la figura 1 resumen los principales resultados del impacto económico asociado a cada universidad, totales y relativizados respecto al número de estudiantes y de ocupados en la universidad, para poder compararlos. El cuadro 4 presenta los porcentajes de PIB de cada región que tiene su origen en cada universidad. Por último, el cuadro 5 muestra la distribución sectorial del impacto.

Considerando el impacto global (como agregación del impacto procedente de los cinco agentes considerados), el cuadro 2 muestra cómo la UCH y UCR son las universidades con más facturación total, superior a los 1.500 MUSD (incluyendo impacto directo, indirecto e inducido), resultado vinculado a la gran dimensión de dichas universidades (con más de 35.000 estudiantes). Les siguen la UACH, la cual, a pesar de tener una dimensión similar a las dos anteriores (con casi 30.000 estudiantes), tiene una facturación total muy inferior a ellas. También destacan la UTP, UNACH y la CES, la cual, con poco más de 6.000 estudiantes, factura casi 184 MUSD, cifra similar a la de UTA o CUC (si bien estas dos últimas universidades tienen un número de estudiantes superior). En el caso de la CES, el importante peso de las actividades relacionadas con la consultoría y los servicios de I+D+I antes comentados explicarían este resultado. El menor volumen de facturación total lo aporta UNNOBA, debido a su reducida dimensión, siendo la que menos ocupados tiene de las nueve universidades y la segunda con menor número de estudiantes.

A su vez, en términos de ocupación, se observa cómo las universidades más grandes (UCH y UCR) generan en total casi 30.000 ocupados cada una, seguidas de UACH (con casi 12.000). Resaltar nuevamente el caso de CES que, siendo la universidad con menor número de estudiantes, acaba generando casi 7.500 ocupados globales (cuarta universidad en términos de generación de ocupación global).

El cuadro 2 también permite conocer qué agentes son los que generan mayor actividad económica en cada universidad. La actividad propia de la universidad es la mayoritaria, entre el 40,4% de la CUC y el 82,8% de UCH. En siete de las nueve universidades el porcentaje es igual o superior al 70%. El segundo agente que mayor aportación realiza a la actividad económica es la de "otra actividad empresarial" (que incluye actividades de prácticas empresariales por parte de estudiantes, actividades de voluntariado en entidades, así como actividad empresarial asociada a concesiones en reprografía, restauración, prestación de otros servicios, etc). La CUC es la universidad con un mayor porcentaje (41,3%), seguido de la UACH (27,2%). También destacan la CES (16,0%), UTP (14,9%) y la UNACH (13,5%). De estas cinco, dos (CUC, CES) son privadas, y la UACH, UNACH y CES tienen una elevada relación con el sector

[4] A partir de ahora, se utilizará el concepto genérico de "región" para aludir a la desagregación territorial inferior al país utilizada en cada caso.

empresarial, vía parque tecnológico (o similar) y actividades de prestación de servicios empresariales de transferencia de conocimiento (vía Uniones Temporales de Empresa, por ejemplo).

En general, el porcentaje aportado por entidades vinculadas a la universidad, pero que tienen CIF (Código de Identificación Fiscal) propio, es muy bajo, casi insignificante. Solo dos universidades superan el 1% (UTA, y UTP). Lo mismo sucede con la generación de actividad proveniente de los visitantes a la universidad (ya sean familiares de estudiantes, o personal académico o administrativo, o visitantes que imparten seminarios, o similares). Sólo en tres universidades (UACH, UCR y UTP) supera el 3%.

Por último, la actividad económica que tiene su origen en los estudiantes (principalmente por desplazamiento a una región distinta de su residencia habitual) es variable, pero mayoritariamente se sitúa alrededor del 10%. A destacar los porcentajes de CUC (17,4%) y UCR (20,9%).

Por último, en el cuadro 2 se relativiza el impacto total global en términos de facturación respecto la facturación directa procedente de las actividades propias de la universidad. Como se puede observar, dicha facturación total global casi duplica, como mínimo, la derivada de la actividad directa de cada universidad. Este es el caso de UNNOBA (ratio igual a 1,9). En el extremo opuesto encontramos a la CUC, para la cual la facturación total global más que quintuplica la facturación directa generada por dicha universidad (acorde con el hecho de ser la universidad donde el peso de la actividad propia de la universidad en el impacto total era el menor entre todas las universidades analizadas). El resto de universidades muestran una ratio entre el 2,2 y el 3,4.

Como complemento a lo anterior, el cuadro 3 permite un análisis comparativo mayor entre universidades, al relativizar las magnitudes respecto al número de estudiantes y ocupados. En primer lugar, en el análisis de los multiplicadores[5], los de facturación son relativamente similares entre universidades (alrededor de 1,6 a 2,2), indicando que la producción directa de la universidad se incrementa entre un 60% y un 120% al considerar los impactos indirectos e inducidos. Algo similar sucede en términos de VAB, con multiplicadores dos décimas inferiores a los anteriores de facturación (y porcentajes entre el 40% y 110%). Los multiplicadores del número de ocupados son similares a los de facturación, excepto en CES (multiplicador mucho más elevado en el ámbito de la ocupación, de 3)[6] y UNACH (multiplicador menor, de 1,4). En general, las universidades colombianas tienen unos multiplicadores superiores al resto de universidades analizadas.[7]

En relación al impacto en facturación, PIB y ocupados generado por las universidades, relativizado al número de estudiantes y ocupados de cada universidad (cuadro 3), en primer lugar, se observa que la ratio estudiante/ocupado ETC de las universidades presentan valores muy diferenciados. Como situaciones extremas se detecta a la CUC (con una ratio de 15,1, resultado del muy escaso número de ocupados para el número de estudiantes que tiene) o a UNNOBA, UNACH o UTA frente a UCH (3,1) y CES (4,3)[8]. Como se puede observar destacan, con diferencia, UCH y CES y esto es algo que se observará con posterioridad en otros indicadores. En gran medida, estos resultados pueden estar originados en las actividades complementarias (antes mencionadas) a las estrictamente docentes en estas dos universidades (por ejemplo, la prestación de otros servicios de consultoría o de atención sanitaria).

[5] Los multiplicadores se calculan como la ratio entre el impacto total global (directo, indirecto e inducido) y el impacto directo global.

[6] Quizá asociado a la especificidad antes mencionada de la Universidad CES, en la que existe una elevada prestación de servicios de transferencia de conocimiento y consultoría. Asimismo, la baja productividad laboral en Colombia (medida por el PIB por ocupado), en comparación a la del resto de países de estudio, podría explicar un rasgo más intensivo en mano de obra de este país que genere un mayor efecto multiplicador en términos de ocupación.

[7] Kureski y Rolim (2009) estiman el impacto a corto plazo de las universidades federales brasileñas. Si bien utilizan la metodologia input-output, los agentes considerados y origen de la información difiere de manera importante de la estrategia seguida en este artículo. A pesar de ello, cabe notar como el multiplicador de empleo que obtienen (3,15) es similar al obtenido para la Universidad CES de Colombia, pero superior al obtenido por el resto de universidades analizadas (único elemento comparable entre su estudio y el desarrollado en este trabajo). Comparando los multiplicadores en las universidades latinoamericanas con respecto a los presentados en los estudios de Suriñach et al (2017, 2023) referidos al sistema de universidades públicas en Cataluña y España (utilizando la misma metodología e hipótesis de trabajo), se obtienen unos multiplicadores algo superiores en las universidades españolas: valores del multiplicador de facturación de 1,8 y 2,4, respectivamente, multiplicador de PIB de 1,6 y 2,2, y multiplicador de la ocupación de 1,6 y 2.

[8] En el caso del sistema universitario público español (SUPE), la ratio es 7,02 (y sin considerar las universidades a distancia es de 6,29).

La facturación total global por estudiante presenta valores dispares entre universidades. Destacan los elevados guarismos de UCH (46.734 USD), UCR (33.028) y CES (29.642). En el extremo opuesto se sitúan CUC (13.664),[9] UACH (14.335) y UNNOBA (9.150). Si se relativiza la facturación total global respecto al número de ocupados ETC de cada universidad, en cambio, el ranking varía (ver figura 1 para una comparativa gráfica de ambas magnitudes). En este caso, las universidades con mejores resultados son UCR (204.670 dólares USD), UTA (176.920), UNACH (131.533), y UCH (145.719)[10].

[9] Cabe recordar que la CUC muestra la ratio de estudiantes por ocupados ETC más elevada de todas las universidades analizadas.
[10] Las universidades que varían más de posición relativa, al considerar facturación total por estudiante o por ocupado, son la CUC (que en el caso de relativizar por el número de estudiantes se sitúa en sexta posición, de las nueve universidades, mientras que en el caso de dividir por ocupados pasa a ser la primera) y la UACH (que pasa de la quinta a la novena, respectivamente).

CUADRO 2.
Análisis comparativo entre universidades. Principales resultados

RESULTADOS VIA DEMANDA (AÑO 2019)	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
Facturación directa global (MUSD)	84,84	86,58	230,21	831,9	888,21	164,85	38,74	80,69	181,57
Facturación total global (MUSD)	183,93	178,73	418,3	1.741,49	1.555,70	286,35	63,16	154,84	312,04
Ocupados directos globales	2.517	2.974	8.272	14.356	14.525	3.735	828	1.181	4.447
Ocupados totales globales	7.428	6.620	11.775	28.401	29.280	6.633	1.302	2.337	7.192
VAB directo global (MUSD)	50,38	50,22	165,4	543,66	687,53	142,91	31,88	60,63	141,88
VAB total global (MUSD)	107,02	101,68	280,76	1.041,82	1.068,57	217,13	44,43	101,38	219,69
Peso de la facturación total derivada de las actividades propias de la Universidad respecto facturación total global	76,8%	40,4%	55,9%	82,8%	71,3%	79,8%	82,4%	80,3%	69,9%
Peso de la facturación total derivada de las empresas vinculadas a la Universidad respecto facturación total global	0,4%	0,0%	0,1%	0,7%	0,0%	0,0%	0,4%	2,2%	3,6%
Peso de la facturación total derivada de la actividad empresarial imputable a la existencia de la Universidad respecto facturación total global	16,0%	41,3%	27,2%	2,8%	4,4%	13,5%	3,9%	2,9%	14,9%
Peso de la facturación total derivada del los visitantes que se desplazan a la ciudad donde se ubica la Universidad respecto facturación total global	0,4%	0,9%	6,0%	0,1%	3,4%	1,8%	1,1%	0,5%	3,1%
Peso de la facturación total derivada del gasto realizado por los estudiantes de la Universidad (MUSD) respecto facturación total global	6,4%	17,4%	10,9%	13,7%	20,9%	4,8%	12,1%	14,1%	8,6%
Facturación total global dividida por facturación directa derivada de las actividades propias de la Universidad	2,9	5,3	3,4	2,6	2,4	2,2	1,9	2,4	2,4

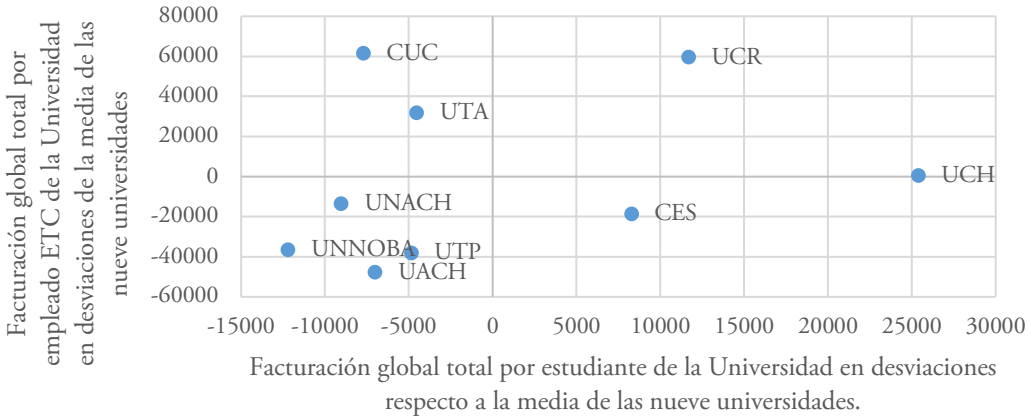
Nota: (MUSD): Millones de dólares de Estados Unidos. Los resultados "totales" surgen de la agregación de los impactos directo, indirecto e inducido. Los resultados "globales" son el resultado de la agregación del impacto procedente de todos los agentes considerados en el estudio.
Fuente: Elaboración propia.

CUADRO 3.
Análisis comparativo entre universidades. Principales Indicadores (relativizados)

	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
DATOS DE LA UNIVERSIDAD (AÑO 2019)									
Estudiantes 2019/Oupados ETC de la Universidad	4,3	15,1	6,8	3,1	6,2	10,7	11,9	10,5	6,5
RESULTADOS VIA DEMANDA (Año 2019)									
Multiplicador facturación	2,2	2,1	1,8	2,1	1,8	1,7	1,6	1,9	1,7
Multiplicador ocupados	3	2,2	1,4	2	2	1,8	1,6	2	1,6
Multiplicador VAB	2,1	2	1,7	1,9	1,6	1,5	1,4	1,7	1,5
Facturación total global por estudiante de la Universidad (USD por estudiante)	29.642	13.664	14.335	46.734	33.028	12.328	9.150	16.838	16.500
VAB total global por estudiante de la Universidad (USD por estudiante)	17.248	7.773	9.622	27.958	22.686	9.348	6.437	11.024	11.617
Ocupados totales globales por cada 100 estudiantes de la Universidad	120	51	40	76	62	29	19	25	38
Facturación total global por ocupado ETC de la Universidad (USD por ocupado)	126.672	206.628	97.331	145.719	204.670	131.533	108.515	176.920	107.010
VAB total global por ocupado ETC de la Universidad (USD por ocupado)	73.708	117.545	65.327	87.174	140.582	99.740	76.343	115.836	75.341
Ocupados totales globales por cada 100 ocupados ETC de la Universidad	512	765	274	238	385	305	224	267	247
Retorno de la Universidad (Ingresos fiscales vía demanda/ingresos procedentes de financiación pública)			63,6%	180,8%	59,8%	31,7%	105,9%	54,5%	29,7%

Fuente: Elaboración propia. Nota 1: (USD): Dólares de Estados Unidos. Los resultados "totales" surgen de la agregación de los impactos directo, indirecto e inducido Los resultados definidos como "global" son el resultado de la agregación del impacto procedente de todos los agentes considerados en el estudio. El multiplicador se computa como la ratio entre el impacto total global dividido por el impacto directo global. Nota 2: En el caso de CES y CUC, no se calcula el "Retorno de la Universidad" dado que son universidades privadas y no reciben financiación pública.

FIGURA 1.
Comparativa de la facturación global por estudiante y por ocupado ETC en desviaciones respecto la media



Nota: Magnitudes monetarias expresadas en dólares de Estados Unidos (USD).
Fuente: Elaboración propia.

Para la variable VAB, los resultados obtenidos anteriormente se mantienen en gran medida (por ejemplo, destacando UCH, CES y UCR como universidades con mayor ratio VAB/estudiante, y UCR, CUC y UTA como universidades con mayor ratio VAB/ocupado).

En relación a las universidades que generan más puestos de trabajo en el país respectivo, relativizando por el número de estudiantes o por el número de puestos de trabajo directo, se obtiene que la CES es la que mejor situada está (creando 120 puestos de trabajo en Colombia por cada 100 estudiantes matriculados y 512 puestos de trabajo por cada 100 ocupados ETC de la universidad). Merece destacarse también a UCR. Respecto a los estudiantes matriculados, destaca también UCH; en cambio, respecto al número de ocupados, destaca la CUC (765 puestos de trabajo por cada 100 ocupados ETC de la CUC). Por último, las universidades que generan menos puestos de trabajo en la economía son UNNOBA (por estudiante y ocupado de la universidad), UTA y UNACH (por estudiante) y UCH y UTP (por ocupado de la universidad).

La parte final del cuadro 3 presenta el porcentaje de retorno (asociado a los ingresos fiscales que genera la actividad universitaria) en relación a la financiación pública recibida por las universidades. A pesar de estar subestimando la actividad (al solo considerarse el impacto a corto plazo), los porcentajes se mueven entre el 30% de retorno (UTP y UNACH) y el 60% (UACH, UCR y UTA), mientras que en dos de ellas supera el 100% (UCH y UNNOBA). Por tanto, el retorno de ingresos al sector público es elevado y, en ocasiones, supera lo aportado directamente por el mismo.

En el cuadro 4 se presenta el porcentaje que representa el PIB generado por la universidad (de manera directa, indirecta e inducida) en el conjunto del PIB del país. Los porcentajes son variables y van del casi 2% en UCR, hasta los valores más pequeños de UACH (0,023%), UNACH (0,017%) o el valor mínimo del 0,012% de UNNOBA. Lógicamente, este resultado está relacionado con la cuota de mercado que representa cada universidad respecto al total de graduados del país (última fila del cuadro 1), apreciándose como UCR concentra el 36% de todos los graduados del país, UTP el 11,7% o UCH el 6,5% frente al 0,2% de UNNOBA.

El cuadro 4 también permite comparar el peso que representa el PIB generado por cada universidad respecto al PIB regional donde se ubica la misma. Nuevamente su peso es variable, dependiendo de la cuota de mercado de la universidad en cada región y del tamaño y actividad económica de la región. En general, la ponderación es de menos del 1%, si bien destacan los porcentajes de la Región Central en el caso de UCR (2,3% de su PIB, con una concentración del 88% de todos sus estudiantes) o de Arica en el caso de UTA (4,1% de su PIB, con una concentración del 84% de sus estudiantes). Ello, sin embargo, no debe

ocultar los importantes valores absolutos de facturación, PIB, rentas salariales, puestos de trabajo y rentas fiscales generadas por cada universidad en su territorio[11].

Finalmente, el cuadro 5 recoge los sectores que concentran el mayor impacto total global en términos de facturación. El sector de Educación es el que concentra, como era de esperar, el mayor impacto en términos de facturación total global. A pesar de ello, dicho impacto no supera, en ningún caso, el 49%, de manera que más de la mitad de todo el impacto en términos de facturación se distribuye por el resto de sectores de la economía. Este hecho demuestra los notables efectos multiplicadores de la actividad universitaria. Entre los sectores más beneficiados, además de la educación, encontramos el sector de actividades financieras y de seguros; comercio; otros servicios del sector empresarial; productos alimenticios, bebidas y tabaco; servicios de alojamiento y alimentación; servicios de suministros energéticos y transporte y almacenamiento.

A partir del conjunto de resultados obtenidos, y a modo de resumen, se destacan las siguientes características de las universidades. En primer lugar, la UCH y, sobre todo, UCR son las dos universidades mejor posicionadas, con ratios elevados de facturación total global tanto por estudiante como por ocupado (ETC) y con mayores impactos en términos de PIB (siendo las dos universidades más grandes tanto en términos de facturación global como de número de estudiantes y de ocupados ETC y con unas elevadas cuotas de mercado en término de los graduados totales del país).

En la situación contraria encontramos a UNACH, UACH y, especialmente, a UNNOBA. Las dos universidades mexicanas, a pesar de tener muchos estudiantes (aunque con pequeñas cuotas de mercado en términos de graduados del país), no tienen facturaciones proporcionalmente más elevadas. Eso lleva a que ambas universidades tengan unas ratios de facturación por estudiante y por ocupados relativamente bajas y con un peso en el PIB mexicano muy reducido (de esperar dada su pequeña cuota de mercado y la dimensión del país). En el caso de la universidad argentina de UNNOBA, se concluye que es una universidad pequeña, tanto en términos de estudiantes como de ocupados (ETC), pero también con una facturación proporcionalmente menor de lo que debería de tener.

Caso diferente es el de CES. A pesar de ser una de las universidades del estudio con menos estudiantes, presenta una facturación proporcionalmente mayor de lo que sería de esperar por su dimensión (la importancia de las actividades de consultoría y de I+D+I explicarían este resultado) y un número de ocupados por estudiante también mayor. Este hecho explica que salga muy bien posicionada cuando las magnitudes de facturación o VAB se relativizan por estudiantes, pero no así cuando se relativizan por ocupados ETC de la universidad.

La CUC y UTA son dos universidades de dimensiones relativamente pequeñas tanto en términos de estudiantes como de facturación, pero con un número de ocupados proporcionalmente más pequeño de lo que sería de esperar por el número de estudiantes. Por ello, presentan valores relativamente bajos de facturación por estudiante, pero elevados cuando se relativiza por el número de ocupados ETC de las mismas.

Por último, la UTP es una universidad de dimensión media en términos de estudiantes y ocupados, pero con una facturación total generada no relativamente elevada, al mostrar unas ratios de facturación por estudiante y por ocupado relativamente bajas. En términos de impacto en el PIB del país, mostraría un valor similar al de, por ejemplo, UCH, hecho que se vería explicado en parte por la relativamente elevada cuota de mercado que muestra en relación a los graduados totales del país.

11 A modo de comparación, a pesar de las enormes diferencias entre economías, Suriñach et al (2017, 2023) empleando la misma metodología, obtuvo un impacto de todo el sistema universitario público de Cataluña del 1,4% sobre el PIB de esta región española o del 2,27% del PIB español en el caso del Sistema Universitario Público Español. Pastor et al (2010a, 2010b) obtienen un impacto del 1,97% del PIB regional en el caso de la Universidad de Cantabria o del 0,91% del PIB regional en el caso de la Universidad de las Islas Baleares en España. Suriñach et al (2021) obtienen un impacto del 0,3% del PIB regional de Cataluña para el caso de la Universidad Politécnica de Cataluña (UPC).

6. Conclusiones

El análisis realizado confirma el elevado impacto económico que generan las universidades latinoamericanas seleccionadas sobre su territorio en el corto plazo. Ello se debe a varios factores, entre los que destacan los dos siguientes. Primero, que el impacto no se circunscribe a un impacto directo (que captaría la facturación generada por cada universidad o los ocupados que trabajan en la misma) sino que se generan impactos indirectos e inducidos adicionales. De esta manera, se generan unos efectos multiplicadores que, en función de la variable, oscilan entre valores de 1,4 a 3.

Dichos efectos multiplicadores, además, afectan a todos los sectores económicos de todos los países implicados. Así, menos del 50% del impacto total se concentra en el sector educativo, viéndose beneficiados el resto de sectores (especialmente, sectores como el comercio, transportes y almacenamiento, actividades financieras, productos alimentarios u otros servicios a las empresas).

En segundo lugar, el impacto universitario no se genera únicamente como consecuencia de la actividad de la propia institución universitaria, sino que existen otros agentes, íntimamente vinculados con ésta, que generan un claro efecto económico que no ese produciría sin dicha institución. Nos referimos a los estudiantes universitarios, a los visitantes de las universidades que realizan un gasto en la región/país donde está ubicada la universidad (ya sea por la asistencia a congresos, a jornadas o por las visitas a familiares), a los Parques Científicos y Tecnológicos u otras instituciones/empresas como fundaciones o consorcios y a todas aquellas empresas cuya actividad tiene su origen en la existencia de la universidad (empresas ubicadas en los Parques Científicos y Tecnológicos, las *start-ups* y *spin-offs* originadas, empresas adjudicatarias de concesiones administrativas y empresas que acogen estudiantes en prácticas y/o de voluntariado).

La consideración de todos los agentes y todos los impactos generados lleva a que el impacto económico en términos de facturación total global más que duplique la facturación directa generada por cada universidad (en algunos casos, como en la CUC, dicho efecto es cinco veces superior).

Cuando el impacto universitario se cuantifica en términos de valor añadido, se obtiene una contribución al PIB regional que oscila entre un mínimo del 0,025% (en el caso de UNNOBA en la provincia de Buenos Aires) y un máximo del 4,1% (en el caso de UTA en la provincia de Arica). Los impactos a nivel nacional son escasos dado que estamos considerando únicamente el impacto de una sola universidad a nivel nacional. A pesar de ello, el impacto sobre el PIB de cada país se movería entre el 0,012% (en el caso de UNNOBA sobre el PIB de Argentina) y un 1,94% (en el caso de UCR sobre el PIB de Costa Rica), valor que varía especialmente en función de la cuota de mercado en términos de graduados totales a nivel nacional de cada una de ellas.

La cuantificación de los ingresos fiscales generados confirma el retorno que tienen los subsidios y ayudas realizadas por la administración a las universidades. Además, los retornos fiscales (ingresos fiscales generados en relación a los recibidos) son elevados, incluso superiores al 100% en algunos casos.

No se detecta que la titularidad pública o privada de la universidad sea el factor más determinante del impacto que genera en la economía. Hay otros factores más relevantes como el de realizar otras actividades de I+D+I o de prestación de otros servicios (que conducen a unos indicadores más competitivos en cuanto a impacto económico generado).

La comparabilidad de estos resultados con los obtenidos en otros estudios de impacto siguiendo un enfoque de demanda es complejo, tanto por lo que se refiere a tipología de impactos estimados (en algunos casos solo se estiman impactos directos e indirectos), como por los agentes generadores de gasto considerados como por las características particulares de cada Universidad y región consideradas.

Del conjunto de resultados obtenidos y del evidente retorno fiscal que generan las universidades, se derivan unas recomendaciones a los gestores de políticas públicas. Entre ellas, merece destacarse, en primer lugar, la idoneidad de legislar para favorecer un modelo universitario más completo, en el que la actividad docente se complemente con otras actividades de investigación, transferencia de conocimientos y prestaciones de servicios a la sociedad (docencia no reglada, consultoría, creación de empresas de base tecnológica, etc). La escasa presencia de parques científicos y tecnológicos, y la aún incipiente creación de

nuevas empresas de base tecnológica, permite abrir el camino para un incremento del impacto generado por las universidades latinoamericanas. Esta recomendación se apoya tanto en el mayor impacto económico obtenido por las universidades con una mayor actividad en investigación y transferencia como en los impactos a largo plazo que la inversión en I+D genera sobre el crecimiento económico (según se ha expuesto en la sección 2). En segundo lugar, parece recomendable intensificar las transferencias en aquellas universidades que ya presentan mejores resultados y sean más eficientes.

A pesar de todos los resultados anteriores obtenidos, cabe remarcar que el estudio adolece de algunas limitaciones que abren la puerta a nuevos análisis. Entre ellas destacan las siguientes. En primer lugar, aparecen las limitaciones asociadas a la propia metodologia input-output para la estimación del impacto entre las que destacan la asunción de proporcionalidad y linealidad al considerarse unos coeficientes técnicos fijos (no se contempla la posibilidad de economías de escala o de cambios en la productividad), el hecho de que no contemplan restricciones de oferta de los factores productivos como el trabajo o, por ejemplo, la no consideración de posibles cambios en los precios o salarios como consecuencia de cambios en la demanda.

La segunda limitación hace referencia a la territorialización de los impactos económicos realizados, especialmente en el caso de los impactos indirectos e inducidos y en el caso de universidades que estén localizadas en más de una región. Dicha territorialización se basa en supuestos que, si bien han tratado de estar respaldados a partir de la información proporcionada por las propias universidades, está sujeta a posibles sesgos o incorrecciones difícilmente contrastables.

La tercera limitación del estudio se centra en una notable subestimación del impacto económico real de las universidades sobre los territorios en los que están ubicadas. Ello se debe a que no se han considerado los efectos más a largo plazo que genera el sistema universitario ("enfoque de oferta" explicado en la sección dos) y que suponen una aceleración del crecimiento económico al incrementar los factores productivos claves de una economía como son la fuerza laboral, el capital humano y el capital tecnológico.

En cuarto lugar, no se ha realizado un estudio detallado del sistema educativo de los países analizados, el cual podría complementar los resultados obtenidos. Por último, tampoco se han considerado los impactos sociales positivos que generan las universidades (Montesi y Villaseñor, 2018; Pastor et al, 2019). Así, el sistema universitario afecta también a otras muchas dimensiones más allá de la economía como es el caso de la cultura, el medio ambiente, la salud, la gobernanza, la alfabetización científica o la cohesión social (atracción de empresas e innovaciones, etc.).

Sin embargo, a pesar de las limitaciones anteriores, debe mencionarse que el estudio presentado supone un avance en el conocimiento del impacto económico generado por las universidades latinoamericanas, por cuanto es el primer estudio que compara nueve universidades de seis países distintos, cuantificando el impacto a partir de una metodología e hipótesis comunes.

CUADRO 4.
Análisis comparativo entre universidades. Impacto sobre el PIB (relativizados)

RESULTADOS VIA DEMANDA (Año 2019)	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
Impacto sobre PIB del país (%)	0,036%	0,035%	0,023%	0,404%	1,941%	0,017%	0,012%	0,039%	0,329%
Impacto sobre del PIB regional donde tiene sede la Universidad (%)	Antioquia: 0,2%	Atlántico: 0,6% (concentra el 98% de estudiantes) y Nariño 0,03% (concentra el 2% de estudiantes)	Chihuahua: 0,55%	Región Metropolitana de Santiago: 0,73%	Región Central: 2,3% (concentra el 88% de los estudiantes)	Chiapas: 0,97%	Província de Buenos Aires: 0,025%	Arica: 4,1% (concentra el 84% de estudiantes) y Tarapacá: 0,2% (concentra el 16% de estudiantes)	Bocas del Toro 0,68% (concentra 0,84% de estudiantes); Chiriqui 0,65% (concentra 10,87% de estudiantes); Cocle 0,53% (concentra 3,99% de estudiantes); Colon 0,06% (concentra 3,06% de estudiantes); Los Santos 1,63% (concentra 5,48% de estudiantes); Panamá Oeste 0,38% (concentra 7,16% de estudiantes); Veraguas 1,24% (concentra 6,33% de estudiantes); Prov. Panamá 0,31% (concentra 62,27% de estudiantes)

Nota: En el caso de UCR, no se muestra el impacto sobre el PIB en el resto de territorios donde existen estudiantes (Guanacaste, Limón y Puntarenas) dado que no se dispone del dato de PIB. En el caso de la Región Central, se ha procedido a estimar el dato de PIB considerando el peso que tiene dicha región en términos de población ocupada a nivel de país.
Fuente: Elaboración propia.

CUADRO 5.
Análisis comparativo entre universidades. Distribución sectorial de la Facturación total global.

	CES	CUC	UACH	UCH	UCR	UNACH	UNNOBA	UTA	UTP
Educación	32,7%	18,7%	28,8%	38,7%	43,1%	48,8%	47,9%	44,8%	40,9%
Actividades financieras y de seguros	4,2%	3,6%	2,0%	2,8%	2,9%	2,3%	3,3%	2,6%	3,1%
Actividades inmobiliarias	3,8%	4,0%	4,3%	4,7%	3,7%	4,5%	2,7%	3,5%	4,0%
Agricultura, silvicultura y pesca	2,8%	2,8%	2,1%	3,1%	2,3%	2,5%	3,0%	2,1%	2,1%
Comercio al por mayor y al por menor; reparación de vehículos de motor	9,6%	7,9%	7,1%	6,3%	5,1%	6,5%	4,3%	5,1%	4,6%
Construcción	3,6%	5,0%	4,4%	2,0%	2,5%	0,6%	5,2%	2,3%	4,3%
Coque y productos petrolíferos refinados	1,3%	1,9%	3,7%	0,8%	0,2%	1,4%	1,3%	0,9%	0,2%
Otros servicios del sector empresarial	5,7%	6,2%	7,4%	14,3%	5,1%	7,3%	4,3%	13,4%	8,1%
Productos alimenticios, bebidas y tabaco	4,2%	3,9%	4,7%	7,2%	5,3%	4,8%	5,7%	3,7%	5,1%
Salud y servicios sociales	10,2%	4,2%	1,4%	2,7%	4,2%	3,6%	1,1%	3,6%	3,3%
Servicios de alojamiento y alimentación	4,8%	6,9%	4,7%	1,2%	8,9%	2,9%	1,2%	4,9%	8,3%
Servicios de electricidad, gas, suministro de agua, residuos y saneamiento	1,9%	17,3%	2,5%	2,1%	2,1%	1,6%	1,0%	1,8%	2,3%
Transporte y almacenamiento	3,6%	5,8%	8,2%	3,2%	4,0%	4,5%	2,5%	3,8%	3,2%
Servicios de TI y otros servicios de información	0,2%	0,7%	0,8%	1,8%	1,5%	0,2%	2,2%	0,2%	0,4%
Resto de sectores	11,4%	11,1%	17,9%	9,1%	9,1%	8,5%	14,3%	7,3%	10,1%

Nota: La distribución sectorial se corresponde con la Facturación total global (agregación de los impactos directo, indirecto e inducido procedente de todos los agentes considerados en el estudio).
Fuente: Elaboración propia.

Agradecimientos

Jordi Suriñach agradece el apoyo financiero brindado por el Ministerio de Ciencia e Innovación para el proyecto titulado 'Desarrollo Humano y Territorial Sostenible', PID2023-146073NB-I00.

Todos los autores agradecen al Banco Interamericano de Desarrollo y a Development Poles el apoyo por el proyecto "Estudio comparativo del impacto económico y social que una selección de instituciones universitarias de América Latina tiene en sus respectivos entornos locales o regionales".

Referencias

Blackwell, M., Cobb, S., & Weinberg, D. (2002). The Economic Impact of Educational Institutions: Issues and Methodology. *Economic Development Quarterly, 16*(1), 88-95. https://doi.org/10.1177/0891242402016001009

Caffrey, J., & Isaacs, H. (1971). *Estimating the impact of a college or university on the local economy*. American Council on Education. https://api.semanticscholar.org/CorpusID:144676319

CEPAL. (2024). *Balance preliminar de las Economías de América Latina y el Caribe, 2024*. https://www.cepal.org/es/publicaciones/81104-balance-preliminar-economias-america-latina-caribe-2024

CEPALSTAT. Base de datos y publicaciones estadísticas, CEPAL. https://statistics.cepal.org/portal/cepalstat/index.html?lang=es

D'Allegro, M., & Paff, L. Q. (2010). Estimating the Economic Impact of Higher Education: A Case Study of the Five Colleges in Berks County. *Pennsylvania Professional File, 117*, 1-17. https://eric.ed.gov/?id=ED511622

Drucker, J., & Goldstein, H. (2007). Assessing the Regional Economic Development Impacts of Universities: A Review of Current Approaches. *International Regional Science Review*, *30*(1), 20-46. https://doi.org/10.1177/0160017606296731

Garrido-Yserte, R., & Gallo-Rivera, M.T. (2010). The impact of the university upon local economy: three methods to estimate demand-side effects. *Annals of Regional Science, 44*, 39-67. https://doi.org/10.1007/s00168-008-0243-x

Hernández, H. H., Martuscelli, J., Moctezuma, D., Muñoz, H., & Narro, J. (2015). Los desafíos de las universidades de América Latina y el Caribe. ¿Qué somos y a dónde vamos? *Perfiles educativos*, *XXXVII* (147), 202-218. https://www.elsevier.es/es-revista-perfiles-educativos-85-articulo-los-desafios-universidades-america-latina-S0185269815000136

Kelly, U., McLellan, D., & McNicoll, I. (2014). *The impact of universities on the UK economy*. Universities UK.

Kureski, R., & Rolim, C. (2009). Impacto econômico de curto prazo das universidades federais na economia brasileira. *Revista Paranaense de Desenvolvimiento*, *117*, 29-51.

Miller, R.E., & Blair, P.D. (2009). *Input-Output Analysis. Foundations and Extensions*. Cambridge University Press. http://digamo.free.fr/io2009.pdf

Montesi, M., & Villaseñor Rodríguez, I. (2018). El impacto social de las instituciones de educación superior: un estudio de caso con la Universidad Complutense de Madrid. *Información, cultura y sociedad: Revista del Instituto de Investigaciones Bibliotecológicas*, (39), 37-60. https://www.redalyc.org/articulo.oa?id=263057241003

Moreno-Brid, J.C., & Ruiz-Nápoles, P. (2009). La educación superior y el Desarrollo económico en América Latina. Serie *Estudios y Perspectivas* (106). CEPAL.

https://repositorio.cepal.org/server/api/core/bitstreams/b7f72ca2-f10d-41d6-ae9a-f43c6350c7b9/content

Nauffal, D. I. (2019). Measuring the economic impact of universities operating un unstable environaments: the case of a private institution in Lebanon. *Higher Education Quarterly*, *79*(73), 343-358. https://doi.org/10.1111/hequ.12200

Pastor, J.M. (dir.), Peraita, C., Pérez, C.A., Soler, Á., & Zaera, I. (2010a). *La contribución socioeconómica de la Universidad de Cantabria.* IVIE. https://www.ivie.es/es_ES/ptproyecto/impacto-socioeconomico-de-la-universidad-de-cantabria/

Pastor, J. M. (dir.), Peraita, C., Pérez, C.A., Soler, Á., & Zaera, I. (2010b). *La contribución socioeconómica de la Universitat de les Illes Balears.* IVIE.

Pastor., J. M., Peraita, C., & Pérez, F. (2016). Estimating the long-term economic impacts of Spanish universities on the national economy. *Papers in Regional Science*, *95*(4), 673-692 https://doi.org/10.1111/pirs.12157

Pastor, J. M. (dir.), Aldás-Manzano, J., Goerlich, F.J., Pérez, P. J., Serrano, L., Catalán, A., Soler, Á., Zaera I., & Mollá S. (2019). *La contribución socioeconómica del sistema universitario español: Informe SUE 2018*. Madrid: CRUE (Conferencia de Rectores de las Universidades Españolas): Conferencia de Consejos Sociales. https://www.ivie.es/es_ES/ptproyecto/la-contribucion-social-economica-las-universidades-espanolas/

Pérez, F., Pastor, J.M., & Peraita, C. (2015) *La Contribución Socioeconómica de las Universidades Públicas Valencianas. Tercer Informe del SUPV 2015*. Ivie.

Ryan, J., & Malgieri, P. (1992). *Economic impact studies in community colleges: The short cut method*. National Council for Resource Development. https://eric.ed.gov/?id=ED469387

Segarra, A. (dir.), Llop, M., Marti, F.P., Piñol, J.M., & Teruel, M. (2012). *El poder de les idees. Impactes de la URV sobre la societat*. Publicacions URV.

Suriñach, J., Murillo, J. & Vayá, E. (2017). Impactes econòmics generats pel SiCUP i pel Sistema Públic d'R+D+i sobre l'economia catalana. Laboratori d'Economia Aplicada AQR-Lab de la Universitat de Barcelona y ACUP (Associació Catalana d'Universitats Públiques). En *Impactes socioeconòmics de les universitats públiques i el sistema públic de recerca de Catalunya*, ACUP. ISBN: 978-84-697-5824-3.

Suriñach, J., García, J.R., Murillo, J., Romaní, J., & Vayá, E. (2021). *Impacte econòmic i retorn social de la Universitat Politècnica de Catalunya (UPC)*. Laboratori AQR-Lab de la Universitat de Barcelona. DL B-20508-2021. https://www.upc.edu/ca/sala-de-premsa/pdfs/informe-impacte-economic-generat-per-upc-al-2019.pdf

Suriñach, J., Vayá, E., & Murillo, J. (2023). Estudio de impacto económico de las universidades públicas españolas. análisis territorializado para el año 2021. Ministerio de Universidades. Gobierno de España. https://www.universidades.gob.es/wp-content/uploads/2023/06/Report_DEF_120623-1.pdf

Tripp Umbach. (2018). *Growing Nebraska. The economic impact of the University of Nebraska* https://nebraska.edu/-/media/unca/docs/university-affairs/state-gov-relations/tripp-umbach-university-of-nebraska-economic-impact-final-study-4-23-2019.pdf?la=en

University of California San Diego. (2018). *Economic, Research, and Community Impacts*. UC San Diego. https://ucsd.edu/_files/UCSD-Economic-Impact-Report-2019.pdf

Valero, A., & Van Reenen, J. (2019). The Economic Impact of Universities: Evidence from Across the Globe. *Economics of Education Review*, *68*, 53-67.

ORCID

Jordi Suriñach https://orcid.org/0000-0003-2622-3280
Esther Vayá https://orcid.org/0000-0001-5661-5627
Jordi López-Tamayo https://orcid.org/0000-0002-5293-6738
Joaquim Murillo https://orcid.org/0000-0002-7213-9437

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 113-128
https://doi.org/10.38191/iirr-jorr.24.061

Articles

Desarrollo local y migración interregional de trabajadores altamente calificados. El caso de Chile

*Rodrigo Pérez Silva**

Recibido: 29 de diciembre de 2023
Aceptado: 05 de agosto de 2024

Resumen:

La capacidad que tienen los territorios locales de crear y atraer capital humano resulta clave en la generación de progreso a nivel local. Sin embargo, los trabajadores calificados tienden no solo a ser formados en determinadas regiones, sino que a sentirse atraídos hacia las mismas. Este estudio analiza los movimientos migratorios interregionales en Chile, identificando los diferenciales de desarrollo económico local como factor que propicia una migración selectiva de trabajadores. Los resultados indican que regiones con altos niveles de desarrollo relativo atraen migrantes más calificados y expulsan a los menos calificados, reforzando así la inicial concentración del capital humano.

Palabras clave: Desarrollo económico local; migración interregional; trabajadores calificados.
Clasificación JEL: R11; R12; R23.

Local development and interregional migration of highly skilled workers. The case of Chile

Abstract:

The ability of local territories to create and attract human capital is key in generating progress at the local level. However, skilled workers not only tend to stem from certain regions, but to be attracted to them. This study analyzes interregional migratory movements in Chile, identifying differentials in local economic development as a factor that promotes selective migration of workers. The results indicate that regions with high levels of relative development attract more qualified migrants and expel the less qualified, thus reinforcing the initial concentration of human capital.

Keywords: Local economic development; interregional migration; skilled workers.
JEL Classification: R11; R12; R23.

1. Introducción

La población y los trabajadores migran, de forma principal, en busca de mejores oportunidades laborales, tales como mejoras salariales y mayores oportunidades de empleo (Dustmann, 2003; Yankow, 1999). A nivel local, esto usualmente implica un movimiento desde áreas rurales hacia centros urbanos (Selod & Shilpi, 2021). Así, no es sorpresa que una proporción importante de la población mundial viva en áreas urbanas (desde el año 2007, más del 50% de los habitantes a nivel mundial viven en áreas urbanas), siendo América Latina y el Caribe la región más urbanizada del mundo (casi el 80% de la población de la

* Departamento de Economía Agraria, Pontificia Universidad Católica de Chile. Núcleo Milenio para el Desarrollo Integral de los Territorios (CEDIT). Chile. raperezs@uc.cl
Autor para correspondencia: raperezs@uc.cl

población de la región vive en áreas urbanas) (CEPAL, 2012). Más aún, de acuerdo a Naciones Unidas,[1] la población ha continuado "urbanizándose" de forma continua, esperándose que para el año 2050 al menos dos tercios de la población mundial viva en áreas urbanas.

Chile no ha estado exento a este fenómeno y si en el año 1950 prácticamente el 40% de la población vivía en áreas rurales, para el año 2000 este porcentaje apenas superaba el 13%, y para 2017 llegaba al 12%.3 Como es de esperar, la explicación para las tendencias que se observan en Chile (y en el resto del mundo) está relacionada con movimientos migratorios desde las regiones y de las áreas rurales hacia Santiago y otras ciudades de mayor tamaño, lugares donde se concentra el empleo y donde los salarios promedio tienden a ser más altos.[2] A modo de ejemplo, entre 2012 y 2017 se produjeron 957.074 movimientos migratorios interregionales al interior de Chile y de ellos, 232.721 (24% del total) se dirigieron hacia la Región Metropolitana (RM), siendo este el principal destino para los emigrantes de 10 regiones y el segundo destino favorito de las otras cinco regiones.[3]

Pese a lo anterior, es importante destacar que, si bien la RM atrae a un alto número de emigrantes nacionales, constituyéndose como el principal foco de atracción nacional, también de ella emigran otro gran número de personas. Es más, el saldo migratorio de la RM refleja que más personas abandonaron la región que las que llegaron a ella entre 2012 y 2017, dejando un saldo neto de migración de -69.357 personas, el saldo negativo más alto en el país. Por contraposición, regiones como Valparaíso, Coquimbo, Maule, La Araucanía y O'Higgins, entre otras, reciben más inmigrantes que las personas que emigran de ellas.[3]

La constatación de los movimientos migratorios entre regiones, sin embargo, por sí sola resulta incompleta para analizar sus efectos sobre el desarrollo interno de las regiones, así como para atender a sus impactos sobre los lugares de origen y destino. Más aún, no permite conocer si la migración que se constata en Chile fue finalmente efectiva en proveer a los migrantes de los beneficios esperados en términos de empleo y mejoras salariales, o bien acceso a servicios y beneficios sociales (Abu Hatab et al., 2022; Bryan & Morten, 2019).

Este artículo se enfoca primeramente en la relación que puede existir entre los niveles de desarrollo local y sus efectos sobre la emigración de trabajadores hacia otras regiones. En tal sentido, los niveles de desarrollo local podrían ser un detonante de la emigración interregional, o bien puede aminorarla y de hecho generar el proceso inverso, de atracción de trabajadores. El segundo elemento trata sobre el efecto que la migración de trabajadores produce sobre el desarrollo local tanto en la región de destino como de aquella que se ha abandonado. En ambos casos esta migración puede modificar, por ejemplo, el stock de capital humano local, alterando así de forma diferenciada la oferta de trabajadores y potencialmente los salarios de dicho grupo, así también como la brecha respecto de los otros grupos de la economía local (Mendola, 2012; Nyberg–Sørensen et al., 2002; Vogler & Rotte, 2000).

Se desprende de lo anterior que un elemento importante del análisis de migración es conocer si ésta se produce de forma homogénea a lo largo de la distribución de habilidades y niveles de educación, o si bien determinados grupos de la población (por ejemplo, los más educados) resultan no solo más propensos a migrar, sino que también existen diferencias en los destinos elegidos (Ahmed & Slrageldin, 1994; Borjas et al., 1992; Fratesi & Percoco, 2014).

Si bien las condicionantes generales para la migración han sido ampliamente estudiadas en la literatura, existen elementos que todavía resultan poco conocidos. En la literatura sobre migración internacional, por ejemplo, es común encontrar que la migración ocurre por diferenciales en tasas de empleo y niveles de ingreso, y que los más calificados son más propensos a migrar (Massey et al., 1993). En tal sentido, la migración entre dos territorios ocurre principalmente como una forma de incrementar el retorno al trabajo y es el reflejo de la búsqueda de mejores condiciones de vida por parte de los migrantes, usualmente de los más calificados (Bodvarsson et al., 2015; Borjas, 1989).

[1] Ver "The 2018 Revision of the World Urbanization Prospects". Disponible en: https://esa.un.org/unpd/wup/.
[2] Un área urbana en Chile se constituye cuando se encuentra un conjunto de viviendas concentradas con más de 2000 habitantes o entre 1001 y 2000, pero donde al menos el 50% de sus habitantes se dedica a actividades secundarias y/o terciarias (INE, 2019).
[3] Ver "Ciudades, Pueblos, Aldeas y Caseríos 2019". Disponible en: https://geoarchivos.ine.cl/File/pub/Cd_Pb_Al_Cs_2019.pdf

A nivel local en tanto, las condiciones teóricas para la migración no son realmente diferentes. Así, la literatura enfocada en migración interna usualmente termina concluyendo que, todo lo demás constante, la migración es más frecuente entre los hombres, solteros, jóvenes y los más educados (González-Leonardo et al., 2020). De hecho, según Bodvarsson & Van den Berg (2013) existe sólo una teoría de migración, tanto interna como internacional: ésta es simplemente una aplicación del modelo de capital humano, donde los migrantes lo hacen básicamente por diferenciales en salarios, lo que bien puede extenderse a mejoras en los niveles de bienestar de una forma más amplia.

Lo que resulta menos estudiado en ambas literaturas, pero especialmente a nivel de migración interna para países en desarrollo, es cómo los niveles de desarrollo local pueden afectar esta propensión a migrar y, más aún, cómo la pueden afectar de forma diferenciada para, por ejemplo, reducir la emigración de los trabajadores más calificados y a su vez promover el desarrollo local.. Esta es precisamente la contribución de este estudio, primero no sólo denotar que existe una relación entre los niveles de desarrollo local y los procesos de migración selectiva, sino que, además, reconociendo que existe una mayor propensión a migrar por parte de los más calificados, denotar que también existen movimientos diferenciados para los trabajadores menos calificados, pero que tienen un destino geográfico distinto.

Más aún, aquellos elementos que resultaban obvios y conocidos, como la importante atracción de la RM para los flujos migratorios, pareciera estar cambiando rápidamente. Regiones en los extremos del país y la misma RM serían ahora aquellas que más expulsarían habitantes, mientras que las regiones más atractivas serían la de Coquimbo y Valparaíso. Si bien, algunos de estos elementos han sido notados de forma relativamente reciente (ver, por ejemplo,Rodríguez Vignoli (2019)) se ha puesto en evidencia que existe una falta de entendimiento profundo sobre el fenómeno y de las causas asociadas a estos cambios. De hecho, el propio Rodríguez Vignoli (2019) señala que hacen falta estudios que logren escudriñar en las causas para estas dinámicas, poniendo especial atención a elementos usualmente ajenos en la discusión, como son aquellos referidos a las amenidades naturales, elementos que afectan la calidad de vida, y la selectividad de los migrantes salientes y entrantes.

Este estudio se enfoca precisamente en describir y caracterizar las dinámicas migratorias regionales, centrándose en los flujos hacia las grandes ciudades, y en la relación que existe entre estos flujos migratorios y los niveles de desarrollo de las regiones y áreas desde donde la población emigra. Finalmente, interesa describir si es que dicho proceso difiere cuando se consideran trabajadores con distintos niveles educativos, potencialmente generando "fugas de cerebros" desde algunas áreas y concentración del capital humano en determinadas ciudades.

Así, el foco de este manuscrito no está puesto en entender los efectos de la migración sobre el bienestar de los propios migrantes, algo ampliamente estudiado en la literatura, sino más bien en entender de qué forma el desarrollo local se relaciona diferenciadamente con la propensión a migrar de distintos grupos de trabajadores, de acuerdo a su nivel de calificación, y en sugerir los efectos que eso puede a su vez tener sobre el mismo desarrollo local, especialmente en zonas rezagadas. En tal sentido, la hipótesis principal del texto es que los diferenciales en niveles de desarrollo local determinan distintos patrones de movilidad según la calificación del individuo.

Finalmente, antes de adentrarse en los resultados, es importante mencionar que este estudio no está exento de limitaciones. Si bien estas son discutidas más adelante, es quizás relevante señalar ahora las dos más relevantes y que pueden ayudar a la comprensión del manuscrito. Primero, la encuesta utilizada no permite seguir a las mismas personas en el tiempo, lo que inhibe de conocer los resultados del proceso migratorio y si es que las condiciones y características de los migrantes (y no migrantes) varían en el tiempo, además de limitar la capacidad de realizar estimaciones potencialmente causales. Segundo, con la información que se cuenta sólo es posible conocer dónde residían las personas 5 años antes del momento de la entrevista. En tal sentido, si una persona migró hace 6 años o más aparecerá clasificada como no migrante en nuestra muestra. Si bien ambos elementos son importantes de tener en cuenta, toda vez que interesa concentrarse en la relación entre niveles de desarrollo y migración relativa, y que tales elementos resultan relativamente estables en el tiempo, estos aspectos no debieran alterar los patrones generales de movimientos migratorios (que es el foco del estudio) ni las estimaciones puntuales en los modelos presentados. No está demás señalar que las estimaciones presentadas deben ser consideradas asociaciones y no impactos causales.

2. Desarrollo local y emigración

Desde el modelo de Harris & Todaro (1970) en adelante, la concepción habitual para los movimientos migratorios es que estos se producen por diferencias salariales observables entre áreas rurales y urbanas, usualmente favoreciendo a las últimas. En tal sentido, un trabajador dado evalúa el salario esperado a obtener en una región distinta a la que reside, y si ese salario resulta más alto, entonces decide emigrar. Modelos posteriores incorporan en este modelamiento no solo la brecha salarial entre dos regiones, sino que además el flujo de ingresos que el trabajador obtendrá a lo largo del tiempo, condicional en la probabilidad de encontrar empleo y el costo de migrar desde una región a otra.

Lo que está implícito en estos modelos es que los trabajadores tienen información sobre algunos elementos observables de los mercados laborales, tales como salarios y tasas de desempleo, y que toman la decisión buscando maximizar su utilidad (Borjas et al., 1992; Korpi & Clark, 2015; Salisbury, 2014). Adicionalmente implícito en los modelos de migración está el hecho de que existe un agente representativo, el que reacciona de determinada forma ante los incentivos de los distintos mercados a los que puede acceder.

Sin embargo, es al menos discutible si este agente representativo es efectivamente representativo de los distintos grupos de habitantes posibles de encontrar en un área determinada. De forma más precisa, si bien es conocido y ampliamente documentado el hecho de que los trabajadores más calificados y más jóvenes, entre otros, son más propensos a migrar (Hunt, 2004), no solo ellos migran y cuando lo hacen, tampoco todos se mueven desde y hacia las mismas áreas (Combes et al., 2012). Hay, en este sentido, una importante heterogeneidad en movimientos migratorios (Liu & Shen, 2017), la que merece atención por cuanto impacta finalmente las posibilidades de desarrollo local de distintas áreas de los países, particularmente de aquellos en desarrollo. Este es precisamente el foco del estudio, que se diferencia de la literatura previa en tanto no se centra en documentar las condiciones individuales que precipitan una mayor propensión a migrar, ni tampoco se focaliza en los impactos que tales procesos migratorios generan en los mercados laborales (de origen o destino) ni en los cambios en las condiciones socioeconómicas de los migrantes. El interés está puesto en la relación entre brechas de desarrollo local (entre regiones) y su asociación con la propensión a migrar de trabajadores calificados versus no calificados.

En específico, interesa evaluar si el nivel de desarrollo local genera impactos diferenciados en los incentivos a emigrar entre distintos grupos de trabajadores. Por ejemplo, si el desarrollo local implica un incremento en los costos de vida de la población, aun cuando eso pueda significar un incremento también en los salarios, es posible que genere mayores incentivos a la emigración de parte de los trabajadores menos calificados, cuyos salarios puedan haberse incrementado relativamente menos. Si este es el caso, entonces ciudades más grandes, como Santiago, por ejemplo, tenderían a expulsar relativamente más a trabajadores menos calificados y a atraer relativamente más a los más calificados. Algo similar puede, entonces, pensarse cierto para la comparación de áreas rurales y urbanas, siendo estas últimas relativamente más atractivas para los trabajadores más calificados.

Lo anterior contrasta con la idea de que los trabajadores se mueven hacia aquellos lugares donde su mano de obra es más valorada, es decir, hacia aquellos lugares donde dicha mano de obra es más escasa. En cambio, si los trabajadores más calificados se mueven hacia áreas más desarrolladas, entonces se produce un proceso de concentración del capital humano y de incremento en las desigualdades territoriales entre las distintas regiones y al interior de las mismas (Fu & Gabriel, 2012; Perez-Silva & Partridge, 2020).

Pese a que existe mucha evidencia en torno a los efectos que tiene la migración sobre los propios migrantes, en términos de empleo e ingresos principalmente, así como en el desarrollo económico de las áreas de las cuales se escapan y de aquellas a las que llegan, el efecto contrario, el del desarrollo sobre la propensión a migrar, no ha sido suficientemente analizado (White & Lindstrom, 2005). Así, es esperable que niveles diferenciales de desarrollo relativo entre las regiones pueden ir promoviendo selectivamente la emigración de distintos grupos a lo largo de la distribución de habilidades de los mismos.

En resumen, la hipótesis central del manuscrito indica que, todo lo demás constante, la propensión a emigrar de un área sea creciente en el nivel de desarrollo del lugar si se es un trabajador menos calificado (si es que el costo de vida, empujado por los salarios de los trabajadores altamente calificados, crece

relativamente más rápido que los salarios para ese grupo), y que lo contrario sea cierto para trabajadores altamente calificados, quienes se benefician de vivir en áreas donde se concentra el capital humano. Así, si las áreas metropolitanas y urbanas en general, se encuentran actualmente con saldos negativos de migración neta, entonces es posible que lo que estemos observando sea una emigración selectiva de trabajadores con menos calificación hacia ciudades de menor tamaño. Lo contrario, una emigración de trabajadores relativamente más calificados hacia las áreas metropolitanas, se daría también de forma simultánea.

Las disparidades regionales y los incentivos que eso puede generar para la migración interregional al interior de Chile, han sido estudiados previamente por diversos autores. Por ejemplo, Paredes (2013) señala que existe una alta concentración de trabajadores y de capital humano alrededor de la Región Metropolitana, lo que, junto con variaciones en los niveles de recursos naturales, genera importantes disparidades territoriales en los salarios. Algo similar habían ya señalado Aroca & Hewings (2002) sobre los movimientos migratorios entre regiones entre 1977 y 1992, destacando una importante concentración de los migrantes en la RM. De forma más importante, la evidencia presentada para el caso chileno parece señalar que, si bien existen procesos migratorios, estos no logran equiparar los niveles de desarrollo entre las regiones y que para conseguir lo anterior, políticas específicas y efectivas resultarían necesarias (Soto & Torche, 2004).

La literatura previa ha sido enfática en denunciar la excesiva concentración de la población (y del capital humano) alrededor de la RM. Este manuscrito busca aportar evidencia en torno a la migración selectiva de los trabajadores entre regiones y denotar cómo los niveles de desarrollo local pueden afectar estas decisiones. En tal sentido, este documento pone de relevancia el efecto diferenciado que pueden tener los niveles de desarrollo local para atraer o repeler determinados grupos de trabajadores en y desde los territorios.

3. Datos y métodos

Para caracterizar los fenómenos migratorios internos recientes se utiliza la información de la Encuesta de Caracterización Socioeconómica (CASEN) de los años 2000 a 2017. Se define como migrante a aquella persona que señala que hace cinco años vivía en una región distinta de aquella en la que vive ahora. Es importante notar que la encuesta CASEN no pregunta por el momento del movimiento, sino que solamente por la comuna de residencia 5 años atrás. En tal sentido, la data no permite identificar a individuos que hayan migrado hace 6 años o más y, por tanto, ellos no son posibles de clasificar como migrantes. De igual forma, no es posible conocer si la persona inmigró hace un año o bien hace exactamente cinco. Para los efectos de este análisis, se asumirá que la persona o bien migró hace alrededor de cinco años, o bien tomó la decisión de migrar con información observable alrededor de ese período. Bajo cierta estabilidad económica y tendencias relativamente constantes de desarrollo, especialmente entre regiones, lo anterior no debiera constituir un supuesto fuerte.

Sea L_{ijt} la localidad j (regiones 1 a 13)[4] donde vive el individuo i en el período t (2006, 2011, 2017). Si el individuo i es identificado como migrante, quiere decir que en el período $t-5$ (aproximadamente 2000, 2006 y 2011, respectivamente) vivió en una localidad distinta de j. Si la región es j, entonces el individuo es categorizado como no migrante en ese período, aún cuando puede haberse movido internamente en la región.

Para establecer los niveles de desarrollo de la región al momento de la decisión de migración, se construye un índice a partir de un análisis de componentes principales (PCA). En particular, para los años 2001, 2006 y 2011 se toma una serie de variables socioeconómicas, tales como tamaño poblacional, población económicamente activa, porcentaje de trabajadores profesionales y técnicos, porcentaje de trabajadores en servicios, ingresos, y niveles de pobreza, entre otras, y se genera un índice de desarrollo manteniendo el primer componente para cada región y año. Este índice es, por construcción, una medida

[4] Se consideran las 13 regiones en las que se encontraba subdividido el territorio nacional al inicio del período considerado, y no las 16 regiones que actualmente existen en Chile. Estas tres regiones nuevas corresponden a porciones de regiones previas que fueron subdivididas para formar el actual mapa de Chile. La correspondencia de regiones y algunos resultados principales para el año 2017, a modo de ejemplo, se presentan en la Figura 4.

de desarrollo regional relativa al resto de las regiones del país, en la cual la importancia de cada variable considerada en la construcción se determina a través del proceso estadístico de PCA.

Si bien constituye una medida parcial de lo que se puede considerar desarrollo en el sentido amplio y se centra de forma casi exclusiva en lo económico, es un ejercicio similar a lo realizado previamente en la literatura y constituye una técnica simple de agrupar variables y reducir dimensionalidad para formar un constructo como bienestar o desarrollo (ver, por ejemplo, Cahill & Sánchez, 2001; Vyas & Kumaranayake, 2006). La lista completa de las variables consideradas en la construcción del índice de desarrollo y su importancia relativa (pesos relativos) se encuentra descrita en el Cuadro A1, mientras que la caracterización de las variables consideradas se muestra en el Cuadro A2 (ambos en anexos). Cómo se observará en la sección de Resultados, el índice construido guarda una estrecha relación y sortea con éxito el desafío de producir un ranking de desarrollo relativo para las regiones de Chile, acercándose de forma importante a otros rankings que consideran un universo mayor de variables y dimensiones.

Finalmente, para analizar si los niveles de desarrollo relativo impactan la decisión de emigración, se emplea un modelo de regresión logística. En específico, sea y_{ijkt} una variable dicotómica con valor 1 si es que el individuo i ha emigrado desde una región j a otra k (con $j \neq k$) en el tiempo t y 0 si no. El individuo i se asume toma la decisión de migrar en función de una serie de variables observables, tales como características personales (edad, educación, sexo, etc.) y el desarrollo relativo de la región de origen (j). Una primera aproximación al modelo empleado está dada por

$$y_{ijkt} = X'_{it+5}\beta + \delta D_{jt} + \mu_{ijkt+5} \quad (1)$$

Donde X es un vector de características del individuo i (observadas en $t+5$) y D_{jt} es un índice de desarrollo relativo construido en t. Como interesa aproximarse a probar la hipótesis de que existe un efecto diferenciado del nivel de desarrollo local sobre la propensión a migrar de distintos grupos de trabajadores, un segundo modelo incluye un efecto interactivo entre D y el nivel educativo del individuo i. El modelo completo está dado por

$$y_{ijkt} = X'_{it+5}\beta + \varphi S_{it+5} \cdot D_{jt} + L_j + L_k + T_t + \mu_{ijkt+5} \quad (2)$$

La ecuación (2) incorpora efectos fijos de la región de origen (L_j) y una dummy de año, para capturar cambios generalizados en las tendencias migratorias entre años. Los efectos fijos de región producen que la variación provenga exclusivamente de cambios en los niveles de desarrollo relativo al interior de cada región y por tanto las estimaciones no están directamente afectadas por elementos fijos que causan que algunas regiones sean más atractivas a los migrantes que otras.

En la ecuación (2) se está interesado en estimar principalmente el parámetro φ, el que se considera como el efecto diferencial que tiene el desarrollo local relativo sobre la propensión de migración para cada grupo de trabajadores: altamente, medianamente, y menos calificados.

Antes de pasar a los resultados, resulta importante reiterar lo señalado en el último párrafo de la introducción, sobre las limitantes del estudio. La base de datos utilizada no constituye un panel de individuos y por tanto no es posible seguir a las mismas personas en el tiempo (algo que sería relevante de tener para estimar si existen características individuales no observables que afectan la propensión a migrar, más allá de la escolaridad que sí es observable, por ejemplo). Esto no sólo impide el seguimiento de los individuos y la utilización de, por ejemplo, efectos fijos, sino que adicionalmente previene de realizar estimaciones causales y de hablar directamente de impactos o efectos del desarrollo sobre la propensión a migrar. Los resultados que siguen deben ser interpretados, por tanto, como asociaciones y correlaciones entre los niveles de desarrollo y la propensión a migrar, condicional en el nivel educativo del individuo.

De forma similar, sólo se conoce dónde vivía la persona cinco años antes de la entrevista, por tanto, personas que hayan migrado hace 6 años o más, aparecerán clasificadas cómo no migrantes. Sin embargo, y tal cómo se mencionó de forma previa, estos aspectos no debieran alterar los resultados principales, los cuales se enfocan en las condiciones generales de migración, la tasa relativa de migrantes calificados y no calificados, y en su relación con los niveles de desarrollo regional. Toda vez que estos elementos presentan

cierta estabilidad en el tiempo (al menos en términos de las diferencias entre regiones), los patrones generales no debieran verse afectados y por ende las estimaciones son consistentes.

4. Desarrollo regional y migración selectiva ¿Qué dice la evidencia?

Antes de comenzar el análisis econométrico, es importante mostrar las tendencias observadas en términos de desarrollo regional, así también como de movimientos migratorios en el tiempo. El Cuadro A1 (en anexos) presenta los niveles promedio en cada uno de los indicadores incluidos en el índice de desarrollo, para el período 2000-2011.

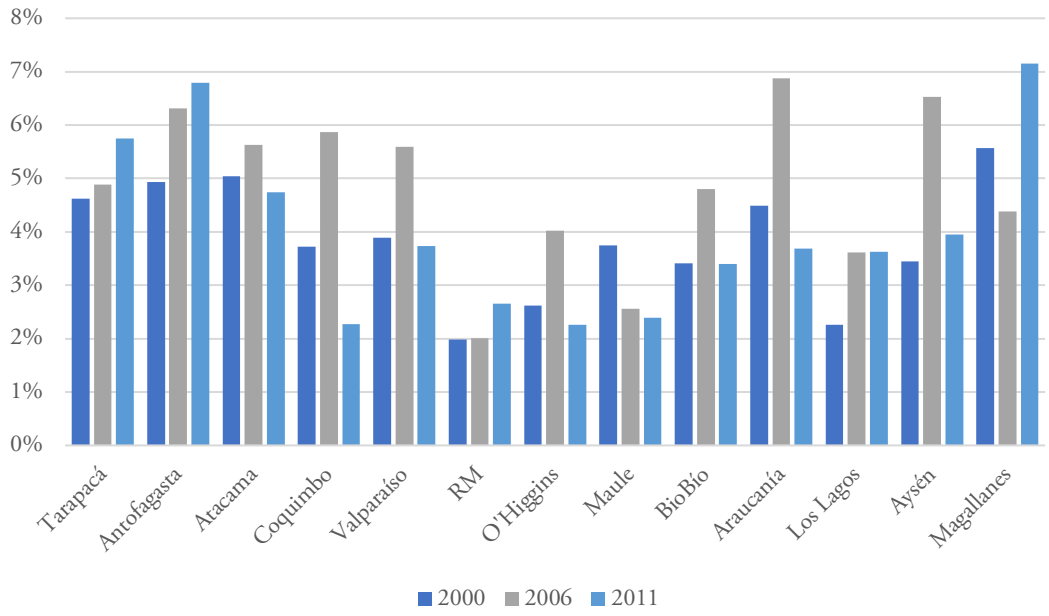
Como se observa en el cuadro A1, los mayores niveles de desarrollo se encuentran en la Región Metropolitana, con un índice que alcanza un valor promedio de 4 para al período considerado, seguido por las regiones de Magallanes (2.6) y Antofagasta (2.4). Por contrapartida, el índice alcanza su valor promedio más bajo en la Región de la Araucanía (-2.5) y del Maule (-2.3). Esto es altamente consistente con lo estimado por el Índice de Desarrollo Regional LATAM realizado por las Universidades Autónoma de Chile y de la República de Uruguay (Cossani et al., 2021). En dicho índice, que considera múltiples dimensiones adicionales de desarrollo, la Región Metropolitana resulta con el nivel de desarrollo más alto del país, junto con Antofagasta. Magallanes es clasificada como medio alto, junto a Valparaíso y Atacama, mientras que la Araucanía, tal como aparece en este estudio, es también considerada en los niveles más bajos de desarrollo (la Región del Maule, en tanto, obtiene una calificación de desarrollo medio bajo). Esta coincidencia entrega amplia confianza en el índice calculado en este estudio.

Lógicamente este resultado es el reflejo de algunas de las características que se observan en cada una de las regiones. A modo de ejemplo, los niveles de pobreza son consistentemente más altos en la zona de la Araucanía, seguida de cerca por las regiones del Biobío y del Maule, donde los promedios 2000-2011 superan el 20% de pobreza. De igual forma, la proporción de trabajadores con títulos profesionales y técnicos, por ejemplo, es más alto en la RM, Magallanes, Antofagasta y Valparaíso, y más bajo en la zona anteriormente descrita. Lo anterior sugiere que los indicadores seleccionados para construir el índice de desarrollo son en su conjunto apropiados para describir disparidades en niveles de desarrollo de forma más amplia que lo que resultaría si solamente se usara un indicador, tal como ingresos o escolaridad, por ejemplo.

En relación con los movimientos migratorios, se observa que la RM es la región que más emigrantes genera y que este proceso solo ha aumentado, y de forma marcada, en los últimos años. Esto es lógicamente consistente con el hecho de que la RM es por lejos el área más poblada de Chile. De igual forma, la siguen las regiones de Valparaíso y Biobío, aun cuando estas dos regiones redujeron el número de personas que emigran para el último año observado. El caso contrario se observa en Tarapacá, Antofagasta y Los Lagos, donde al igual que en la RM, se observa un crecimiento constante en el número de personas que emigran (ver Figuras A1 y A2).

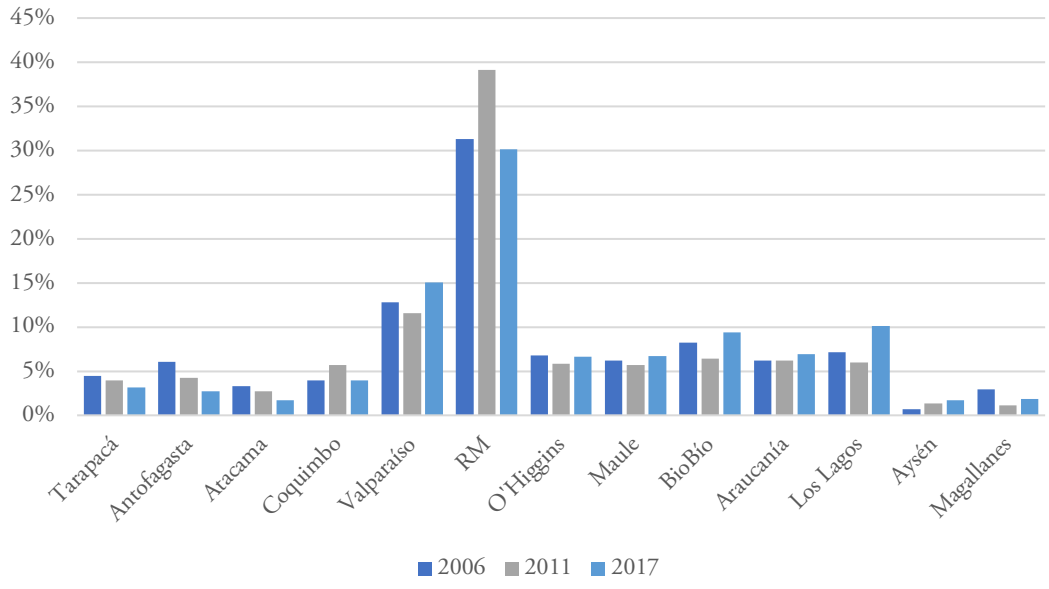
Un análisis más interesante quizás consiste en revisar cómo ha evolucionado la emigración al interior de cada región, pero considerando el número de emigrantes relativo al total de la población de la propia región. La Figura 1 presenta esta información. A modo de ejemplo, la RM es una de las regiones que menos emigrantes genera en todo el período considerado, especialmente si se toman en cuenta los dos primeros años de observación, cuando solo el 2% de sus trabajadores emigraba. Para alrededor del año 2011, este valor alcanzó el 2.7%, por sobre las tasas de emigración de Coquimbo, O'Higgins y Maule. Por contrapartida y centrándose en el año 2011, Magallanes (7.1%), Antofagasta (6.8%) y Tarapacá (5.7%) son las regiones desde las cuales más se emigra y, para el caso de las dos últimas, esta es una tendencia al alza que se mantiene por toda la serie. Resulta interesante notar de igual forma que la región de la Araucanía presentaba hasta el año 2006 uno de los mayores flujos emigratorios del país, con casi 7% de sus trabajadores emigrando hacia otras regiones, tendencia que se revirtió para el año 2011.

FIGURA 1.
Porcentaje de los trabajadores de cada región que emigra, alrededor de los años 2000, 2006 y 2011



Fuente: Elaboración propia en base a CASEN 2006, 2011 y 2017.

FIGURA 2.
Región de destino de los migrantes, alrededor de los años 2006, 2011 y 2017



Fuente: Elaboración propia en base a CASEN 2006, 2011 y 2017.

El lado contrario de la moneda es observar hacia dónde se dirigen los emigrantes. La Figura 2 presenta esta información en función del total de emigrantes de cada año. Nuevamente, la RM es la región que concentra el mayor número y porcentaje de inmigrantes, recibiendo a entre 30 y 40% del total de movimientos interregionales en el período. La segunda región en términos de atracción de migrantes

interregionales es la región de Valparaíso, albergando al 15% del total de inmigrantes del año 2017 y a una proporción similar en los años previos. Del resto de las regiones destacan Los Lagos y Biobío, regiones que reciben alrededor del 10% de los inmigrantes respectivamente, al menos para el año 2017.

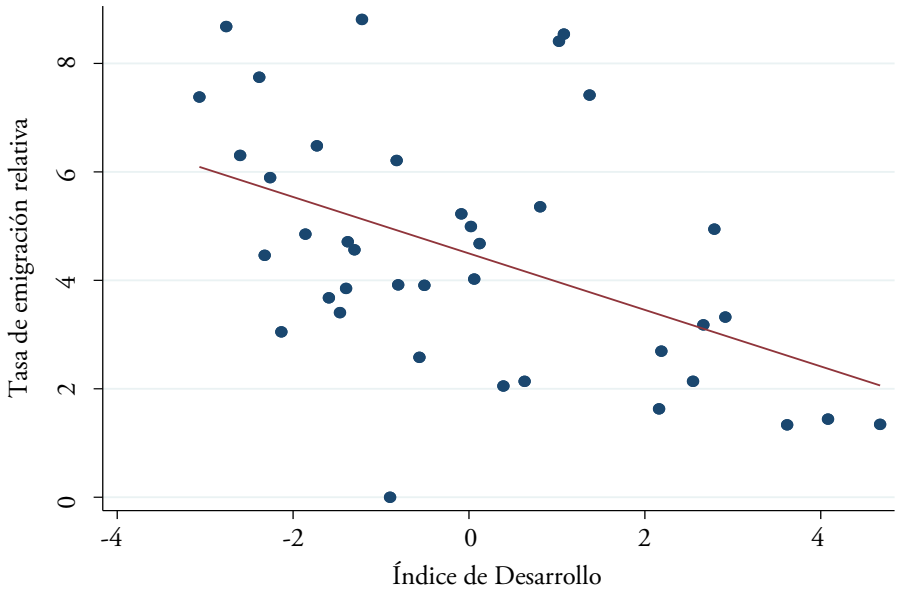
Quizás el elemento más llamativo es la pérdida de atractivo para los inmigrantes de las regiones del norte del país. En las regiones de Tarapacá, Antofagasta y Atacama no solo se presentan tasas bajas de recepción de inmigrantes interregionales, sino que además estos números vienen decayendo de forma constante en el período considerado. Aysén, en el extremo opuesto geográficamente, presenta tendencias contrarias con un incremento (leve) de su atractivo para los migrantes interregionales, aun cuando el porcentaje de trabajadores que emigra hacia esa región es todavía relativamente bajo.

Lo anterior, si bien sugiere una relación entre los niveles de desarrollo y la migración, no permite evidenciar la posible heterogeneidad que pueda existir entre grupos de trabajadores. Una primera aproximación a la pregunta de interés es observar cómo se relacionan los niveles de desarrollo con las tasas relativas de emigración según nivel educativo. Para llevar adelante este ejercicio, se determina la razón entre la tasa de emigración de trabajadores altamente calificados (con título profesional o más) y la de los trabajadores con el menor nivel de calificación en la muestra (menos que Enseñanza Media), y se la compara con los niveles de desarrollo regional para todos los años. Esto es, sea M_s la tasa de emigración de los trabajadores calificados (número de trabajadores calificados que emigran, respecto del total de trabajadores calificados) y M_u la tasa de emigración de los trabajadores no calificados (número de trabajadores no calificados que emigran, respecto del total de trabajadores no calificados), entonces $RMS = M_s/M_u$ es la tasa de emigración selectiva.

La Figura 3 sugiere que los niveles de desarrollo están relacionados con la emigración selectiva de trabajadores calificados. Esto es, a mayor nivel de desarrollo, menor emigración de trabajadores calificados respecto de los no calificados, o, por el contrario, mayor emigración de trabajadores no calificados respecto de los calificados. La Figura 4 muestra esta asociación para el año 2017, con una marcada concentración de las mayores tasas de emigración relativa en la zona central del país (con excepción de la RM de Santiago), lo que se condice con menores índices de desarrollo relativo.

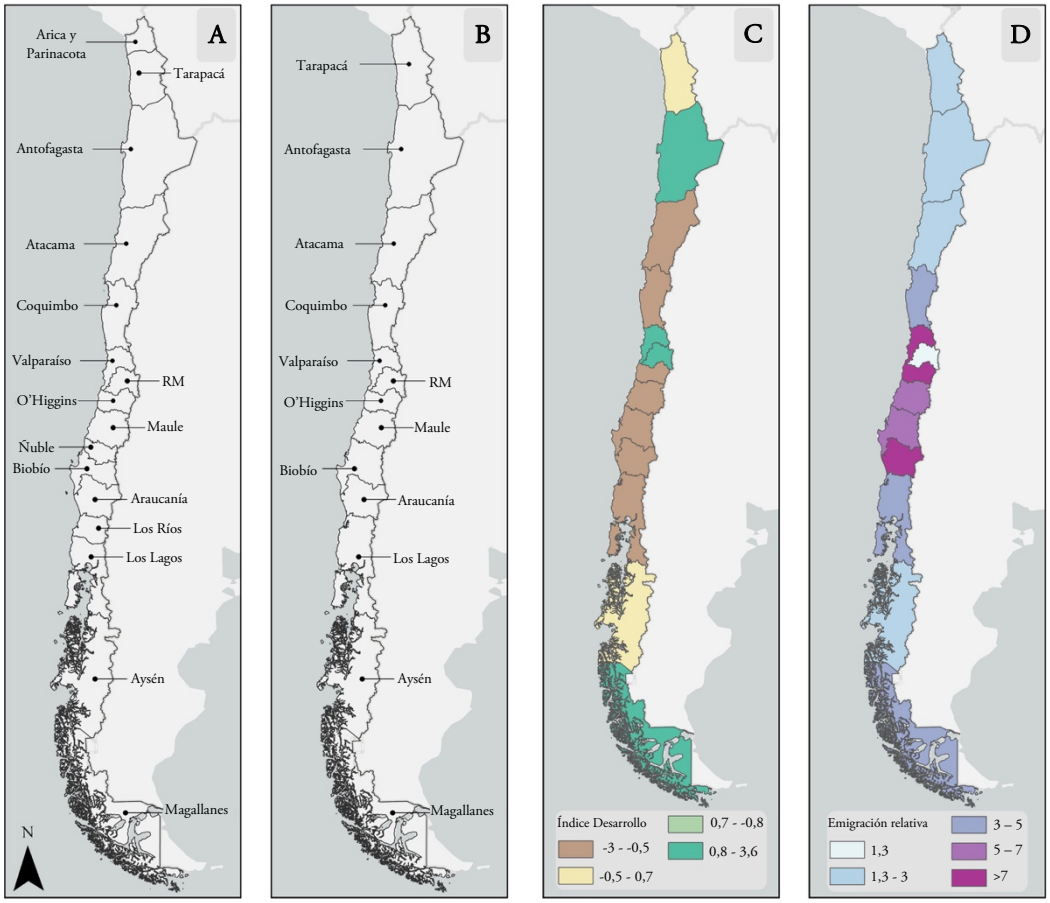
La asociación entre ambas variables, determinada por una regresión simple sin controles, resulta negativa (con un coeficiente de -0.521) y estadísticamente significativa al 1%. Así, mientras mayor el nivel de desarrollo de una región, mayor es la probabilidad de que los trabajadores no calificados emigren de ella. Si bien esto sugiere que la hipótesis inicial es correcta, una comprobación más formal se presenta en los resultados principales.

FIGURA 3.
Tasa de emigración relativa y niveles de desarrollo regional



Fuente: elaboración propia base a CASEN 2006, 2011 y 2017.

FIGURA 4.
Desarrollo regional y tasas de emigración relativa en las regiones de Chile, 2017



Fuente: Elaboración propia en base a CASEN 2017.

5. DESARROLLO REGIONAL Y MIGRACIÓN SELECTIVA. EVIDENCIA EMPÍRICA

En esta sección se presentan y discuten los principales resultados del documento. En el Cuadro 1 se presentan tres modelos, todos con resultados en *odds ratios*. El primero de ellos (columna [1]) considera únicamente las variables de interés más *dummies* de año y de región de origen. El modelo de la columna [2] incluye, además de lo ya mencionado, controles sociodemográficos adicionales descritos en la parte baja del Cuadro. Finalmente, el modelo de la columna [3] repite lo anterior, pero omite del análisis a la RM para evaluar si los efectos considerados son únicamente empujados por dicha región. Cabe notar que, para añadir confiabilidad en los resultados obtenidos, en los tres modelos se agrega el valor del test de bondad de ajuste de Hosmer-Lemeshow y su respectivo valor p. En todos los casos en que el modelo tiene controles (columnas [2] y [3]), el test indica que existe un buen ajuste del modelo logístico. De forma similar, en el Cuadro A3 en anexos se presenta en porcentaje de casos correctamente clasificados en los modelos (P>0.75), los que en todos los casos superan el 96%. Los efectos marginales de las variables consideradas se encuentran en el Cuadro A4, en anexos.

Tomando como punto de partida el modelo sin controles se observa que el índice de desarrollo local se encuentra positivamente correlacionado con la emigración, esto es, cada punto en el índice incrementa la chance de emigrar en un 26%. Como es de esperar, a medida que aumenta el nivel educativo del

individuo, este es más propenso a emigrar. A modo de ejemplo, un trabajador profesional es más de 4 veces más probable que emigre que uno que no terminó la educación media (*odds ratio* de 5.45).

Finalmente, las últimas tres variables antes de la constante interactúan el nivel educativo del individuo con el nivel de desarrollo local para evaluar el efecto diferenciado de esto sobre la propensión a migrar. Los resultados indican, consistentemente con la estadística descriptiva mostrada previamente, que la probabilidad de emigrar se reduce con el desarrollo a medida que el nivel educativo del trabajador aumenta. Esto implica que los trabajadores más calificados tienden a quedarse o a no emigrar de las áreas que tienen un mayor desarrollo local. Alternativamente implica que, consistente con nuestra hipótesis inicial, a medida que aumenta el desarrollo, las áreas tienden a expulsar más a los trabajadores menos calificados. En particular, con el incremento en el desarrollo local, un trabajador altamente calificado tiene un 22% menos de probabilidad de emigrar de un área dada que un trabajador que no completó su enseñanza media.

CUADRO 1.
Asociación entre niveles de desarrollo local y emigración por grupo educativo (*odds ratios*). Modelo *Logit*. Variable dependiente es dicotómica (1 = migra, 0 = no)

Variables	Sin controles [1]	Con controles [2]	Sin RM (Con controles) [3]
Índice Desarrollo	1.258***	1.306***	1.089
	(0.0577)	(0.0622)	(0.0668)
Enseñanza Media	1.847***	1.216***	1.020
	(0.0844)	(0.0807)	(0.0359)
Técnico	3.571***	1.919***	1.497***
	(0.1715)	(0.1831)	(0.1967)
Profesional +	5.451***	2.487***	1.504**
	(0.2636)	(0.3493)	(0.3040)
EM*Índice Desarrollo	0.951***	0.955***	1.020
	(0.0152)	(0.0158)	(0.0359)
Técnico*Índice Desarrollo	0.863***	0.857***	0.951
	(0.0146)	(0.0153)	(0.0349)
Prof + *Índice Desarrollo	0.782***	0.762***	0.885***
	(0.0131)	(0.0135)	(0.0346)
Constante	0.008***	0.338	0.282
	(0.000)	(0.2486)	(0.2629)
Observaciones	122,508	122,508	98,373
Hosmer-Lemeshow chi2 Prob > chi2	16.32 0.038	17.66 0.478	22.01 0.232

Nota: regresiones para jefes de hogar entre 20 y 65 años de edad. Todas las regresiones incluyen *dummies* de año y de región de origen. Las columnas 2 y 3 (Con Controles) además incluyen controles de edad, edad al cuadrado, sexo, número de personas en el hogar, escolaridad, escolaridad al cuadrado, rama de actividad económica, estado civil, condición de pobreza, horas trabajadas y el logaritmo del ingreso total. Errores estándares robustos en paréntesis. El test de bondad de ajuste utiliza 10 grupos en el modelo sin controles y 20 grupos en los modelos con controles. *** p<0.01, ** p<0.05, * p<0.1

De forma interesante, si bien algunos de los coeficientes se reducen de forma importante, los resultados generales se mantienen cuando se incluyen controles en la regresión. El índice de desarrollo pasa de 1.26 en el modelo sin controles a 1.3 en aquel que incluye controles. De igual forma, si anteriormente la probabilidad de emigrar de un trabajador altamente calificado caía 22% respecto a uno no calificado a medida que aumentaba el desarrollo local, esta cifra llega a cerca de 24%, ligeramente por sobre lo estimado por el modelo previo. Quizás el mayor cambio se produce en términos de las diferencias por niveles educativos, los que, si bien siguen sugiriendo que los trabajadores más calificados son más propensos a emigrar, las diferencias se acortan a la mitad respecto al modelo de la columna [1].

Otras variables consideradas, como son el sexo, la edad, el estado civil y otras, presentan los resultados esperados y resultan consistentes con la evidencia internacional previa (ver Cuadro A3 en Anexos). Por ejemplo, las mujeres son cerca de un 40% menos propensas a emigrar que los hombres, mientras que los trabajadores casados emigran casi un 8% menos que los no casados. La edad, por su parte, reduce la propensión a migrar (*odds ratio* de 0.94).

Una preocupación importante que se podría presentar es que, dada la magnitud de la RM, la mayor parte de los resultados estén realmente reflejando solo diferencias entre dicha región y el resto del país. Dicho de otra forma, quizás solo los habitantes más educados de la RM reducen su propensión a migrar cuando la región se desarrolla, pero eso podría no resultar cierto para el resto del país. Si esto fuera así, los coeficientes previamente mostrados deberían desaparecer en significancia en el tercer modelo.

Lo primero que es importante de mencionar con respecto al modelo de la columna [3] es que, cuando se excluye la RM, el efecto parcial del desarrollo parecería desaparecer. De igual forma, el poseer enseñanza media no entregaría una mayor propensión a la migración por sobre aquellos trabajadores que no la completaron. Pese a esto, sigue siendo cierto que niveles educativos superiores, como los de los técnicos y los de los profesionales, resultan significativamente más propensos a emigrar que los menos educados.

Finalmente, al menos para el caso de los más calificados (Profesionales +), el desarrollo local sigue siendo un importante para reducir su propensión a emigrar. Esto es, incluso fuera de la RM, los trabajadores más calificados, se ven inhibidos (más que aquellos con menos que enseñanza media) en su decisión de migración cuando la localidad donde viven se desarrolla. Estas diferencias se mantienen con respecto a todos los otros grupos de forma simultánea. Esto resulta altamente relevante para las políticas públicas, por cuanto permite sostener que variaciones positivas en los niveles de desarrollo local resultan extremadamente relevantes para reducir la "fuga de cerebros" a lo largo de todo Chile y que, por consiguiente, promover el desarrollo local puede tener impactos positivos para generar círculos virtuosos de desarrollo sostenible a través de la creación y atracción de capital humano (ver, por ejemplo, los resultados del estudio de Vidal (1998) para la migración internacional). De esta forma, promover el desarrollo local puede tener efectos significativos para reducir también las disparidades regionales en el largo plazo, generando un proceso de desarrollo autónomo de largo aliento.

La evidencia presentada tanto en esta sección como en la anterior, permiten sostener que el nivel de desarrollo local es en primer lugar importante para condicionar la decisión de emigrar desde un área dada. Segundo, tal propensión si bien aumenta con el nivel educativo (consistentemente con la literatura) se reduce cuando los territorios locales se desarrollan.

Ahora bien, como el índice de desarrollo se ve altamente influenciado por la escolaridad promedio del lugar y por el porcentaje de trabajadores profesionales y técnicos, entre otras variables, lo anterior sugiere que existen ganancias de la concentración del capital humano, las que son generalizables para todo el país y no exclusivas de la RM. En tal sentido, inversiones públicas, por ejemplo, en centros de investigación y universidades son importantes no solo en sí mismos para incentivar el desarrollo, sino que actuarían como un imán para los trabajadores calificados locales, inhibiendo sus intenciones de emigrar, y para aquellos que estén emigrando de otras regiones y áreas del país.

6. Conclusiones

Este estudio buscaba documentar las tendencias de las migraciones interregionales recientes en Chile y relacionar tales flujos con los niveles de desarrollo local regional. Adicionalmente, buscaba evidenciar la existencia de heterogeneidad en los movimientos migratorios, considerando distintos tipos de trabajadores.

Los resultados principales del estudio sugieren que si bien la emigración se ve afectada por las variables sociodemográficas habituales (edad, sexo, estado civil, entre otras), tal como se hipotetizaba, los niveles de desarrollo local juegan un importante rol en promover o reducir la propensión a emigrar. Adicionalmente, dicha propensión a emigrar no actúa de la misma forma entre distintos grupos de trabajadores, empujando más a la emigración a distintos grupos de trabajadores dependiendo del nivel de desarrollo local. Así las áreas más desarrolladas del país, las cuales se extienden más allá de la RM, tienden a concentrar al capital humano y potencialmente a extender estas diferencias interregionales en desarrollo.

Las regiones que más expulsan trabajadores están en las zonas extremas del país, con el norte grande y el área de Aysén y Magallanes como las áreas desde donde más se emigra en términos proporcionales. De igual forma, la mayor parte de la inmigración se concentra en la RM y en las zonas del centro del país. Lo anterior es consistente con el hecho de que los mayores niveles de desarrollo se concentran en la RM, Valparaíso y el Biobío.

Los resultados sugieren que los niveles de desarrollo local son importantes en no solo atraer trabajadores calificados, sino que en reducir la propensión a migrar de aquellos ya localizados en la región. Quizás la única excepción a esta conclusión la puede constituir la región de Valparaíso, donde pareciera producirse un recambio permanente de trabajadores calificados.

Si bien no se han encontrado estudios previos similares al presentado, que busquen simultáneamente vincular el nivel de desarrollo local con la migración selectiva, los resultados encontrados son consistentes con la evidencia internacional, que señala que los trabajadores calificados eligen localizarse en ciudades con mayores oportunidades económicas (mayor desarrollo), y con la lógica de los beneficios de la concentración del capital humano en ciudades, ampliamente discutida en la literatura (ver, por ejemplo, Abel et al., 2012; Fu & Gabriel, 2012; Rauch, 1993; Rotemberg & Saloner, 2000).

Lo anterior resulta relevante, por cuanto una migración selectiva interna produce un círculo virtuoso entre retención de capital humano avanzado en regiones desarrolladas (una menor emigración de los más calificados respecto de los no calificados) y una adicional atracción de capital humano desde regiones más rezagadas. Esto, de acuerdo a la literatura citada, produce un incremento en los niveles de productividad y desarrollo en las mismas regiones donde se concentra el capital humano e incrementa aún más las brechas existentes de desarrollo entre regiones "ganadoras" y aquellas que se quedan rezagadas.

En tal sentido, es posible señalar que estrategias de promoción del desarrollo local pueden tener impactos de largo plazo a través de la retención de los trabajadores calificados, los que a su vez pueden generar desarrollo económico y social sostenido en el tiempo. De esta forma, más allá de los efectos inmediatos que producen las estrategias de desarrollo local, los cuales por cierto resultarán positivos, la reducción en la propensión de emigración de los trabajadores calificados en las economías locales puede ayudar a reducir las disparidades regionales y a mejorar las condiciones de vida en las regiones y áreas rurales en el largo plazo.

Si bien disparidades en los niveles de costo de vida (en favor de las regiones menos avanzadas, que podrían tener costos de vida más bajos) pueden generar incentivos de mercado para la migración de trabajadores calificados en dichos territorios, la falta de oportunidades concretas y la poca profundidad de los mercados laborales locales, puede constituir una barrera difícil de cruzar. En tal sentido, acciones estatales concretas asociadas al fomento al empleo calificado en regiones rezagadas pueden resultar atractivas. Algunas de estas medidas pueden asociarse a la creación de más y mejores programas de postgrado en universidades regionales, que fomenten la creación de masas críticas de académicos y académicas; becas e incentivos de contratación de doctorados y profesionales altamente capacitados en universidades regionales (contra becas de estudio otorgadas por el Estado, por ejemplo), o bien con la condonación o rebaja de la deuda de los estudiantes que estudiaron con crédito en las universidades a

cambio de relocalización en regiones, entre otras. De igual manera, incentivos para que las compañías se relocalicen fuera de las principales regiones también pueden constituir un instrumento relevante para modificar los patrones de migración relativa. Si bien el Estado realiza algunas de estas acciones actualmente, quizás sea necesario profundizarlas y mejorar los incentivos que hoy por hoy no logran producir cambios significativos en las tendencias migratorias de los profesionales chilenos.

Referencias

Abel, J. R., Dey, I., & Gabe, T. M. (2012). Productivity and the Density of Human Capital. *Journal of Regional Science*, *52*(4), 562-586. https://doi.org/10.1111/j.1467-9787.2011.00742.x

Abu Hatab, A., Amuakwa-Mensah, F., & Lagerkvist, C.-J. (2022). Who moves and who gains from internal migration in Egypt? Evidence from two waves of a labor market panel survey. *Habitat International*, 124. https://doi.org/10.1016/j.habitatint.2022.102573

Ahmed, A. M., & Slrageldin, I. (1994). Internal Migration, Earnings, and the Importance of Self-selection. *The Pakistan Development Review*, *33*(3), 211-227.

Aroca, P., & Hewings, G. J. D. (2002). Migration and regional labor market adjustment: Chile 1977-1982 and 1987–1992. *The Annals of Regional Science*, *36*(2), 197-218. https://doi.org/10.1007/s001680200078

Bodvarsson, Ö. B., Simpson, N. B., & Sparber, C. (2015). Migration Theory. In B. R. Chiswick & P. W. Miller (Eds.), *Handbook of the Economics of International Migration* (Vol. 1, pp. 3-1). North-Holland. https://doi.org/10.1016/B978-0-444-53764-5.00001-3

Bodvarsson, Ö. B., & Van den Berg, H. (2013). The Determinants of International Migration: Theory. In Ö. B. Bodvarsson & H. Van den Berg (Eds.), *The Economics of Immigration: Theory and Policy* (pp. 27-57). Springer. https://doi.org/10.1007/978-1-4614-2116-0_2

Borjas, G. J. (1989). Economic Theory and International Migration. *International Migration Review*, *23*(3), 457-485. https://doi.org/10.1177/019791838902300304

Borjas, G. J., Bronars, S.G., & Trejo, S.J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics*, *32*(2), 159-185. https://doi.org/10.1016/0094-1190(92)90003-4

Bryan, G., & Morten, M. (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, *127*(5), 2229-2268. https://doi.org/10.1086/701810

Cahill, M. B., & Sánchez, N. (2001). Using principal components to produce an economic and social development index: An application to Latin America and the U.S. *Atlantic Economic Journal*, *29*(3), 311-329. https://doi.org/10.1007/BF02300552

CEPAL. (2012). *Población, territorio y desarrollo sostenible. Comisión Económica para América Latina y el Caribe*. https://www.cepal.org/es/publicaciones/22425-poblacion-territorio-desarrollo-sostenible

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2012). Sorting and local wage and skill distributions in France. *Regional Science and Urban Economics*, *42*(6), 913-930. https://doi.org/10.1016/j.regsciurbeco.2012.11.003

Cossani, M. C. V., Miranda, A. R., Gorenstein, S., Yanes, P., Sanhueza, P., Rodríguez, I., Parrao, A., Costamagna, P., Rozzi, S. P., Sessa, C., Polenta, C., Sotuyo, M., Castro, R. L., Damiano, S. G., Grin, E. J., Santos, F. B. P. D., Fernandes, G. A. D. A. L., Bresciani, L. P., Neto, L. A., … Rodríguez, C. M. (2021). Medición y agenda para el desarrollo territorial en América Latina. In *Ediciones Universidad Autónoma de Chile*. https://doi.org/10.32457/UA.84

Dustmann, C. (2003). Return migration, wage differentials, and the optimal migration duration. *European Economic Review*, *47*(2), 353-369. https://doi.org/10.1016/S0014-2921(01)00184-2

Fratesi, U., & Percoco, M. (2014). Selective Migration, Regional Growth and Convergence: Evidence from Italy. *Regional Studies*, *48*(10), 1650-1668. https://doi.org/10.1080/00343404.2013.843162

Fu, Y., & Gabriel, S. A. (2012). Labor migration, human capital agglomeration and regional development in China. *Regional Science and Urban Economics*, *42*(3), 473-484. https://doi.org/10.1016/j.regsciurbeco.2011.08.006

Harris, J. R., & Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, *60*(1), 126-142.

Hunt, J. (2004). Innis Lecture: Are Migrants More Skilled than Non-Migrants? Repeat, Return, and Same-Employer Migrants. *The Canadian Journal of Economics / Revue Canadienne d'Economique*, *37*(4), 830-849.

Korpi, M., & Clark, W. A. V. (2015). Internal migration and human capital theory: To what extent is it selective? *Economics Letters*, *136*, 31-34. https://doi.org/10.1016/j.econlet.2015.08.016

Liu, Y., & Shen, J. (2017). Modelling Skilled and Less-Skilled Interregional Migrations in China, 2000-2005. *Population, Space and Place*, *23*(4), e2027. https://doi.org/10.1002/psp.2027

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review*, *19*(3), 431-466. https://doi.org/10.2307/2938462

Mendola, M. (2012). Rural out-migration and economic development at origin: A review of the evidence. *Journal of International Development*, *24*(1), 102-122. https://doi.org/10.1002/jid.1684

Nyberg–Sørensen, N., Hear, N. V., & Engberg–Pedersen, P. (2002). The Migration-Development Nexus Evidence and Policy Options State-of-the-Art Overview. *International Migration*, *40*(5), 3-47. https://doi.org/10.1111/1468-2435.00210

Paredes, D. (2013). The Role of Human Capital, Market Potential and Natural Amenities in Understanding Spatial Wage Disparities in Chile. *Spatial Economic Analysis*, *8*(2), 154-175. https://doi.org/10.1080/17421772.2013.774094

Perez-Silva, R., & Partridge, M. D. (2020). Concentration of human capital, externalities and the wage gap in US metro areas. *Regional Studies*, *54*(11), 1564-1573. https://doi.org/10.1080/00343404.2020.1733515

Rauch, J. E. (1993). Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics*, *34*(3), 380-400. https://doi.org/10.1006/juec.1993.1042

Rodríguez Vignoli, J. (2019). *Migraciones internas en Chile, 1977-2017: Continuidad y cambio*. https://hdl.handle.net/11362/45047

Rotemberg, J. J., & Saloner, G. (2000). Competition and human capital accumulation: A theory of interregional specialization and trade. *Regional Science and Urban Economics*, *30*(4), 373-404. https://doi.org/10.1016/S0166-0462(99)00044-7

Salisbury, L. (2014). Selective migration, wages, and occupational mobility in nineteenth century America. *Explorations in Economic History*, *53*, 40-63. https://doi.org/10.1016/j.eeh.2014.02.001

Selod, H., & Shilpi, F. (2021). Rural-urban migration in developing countries: Lessons from the literature. *Regional Science and Urban Economics*, *91*. https://doi.org/10.1016/j.regsciurbeco.2021.103713

Soto, R., & Torche, A. (2004). Spatial inequality, migration and economic growth in Chile. *Cuadernos de Economía*, *41*(124), 401-424. https://doi.org/10.4067/S0717-68212004012400005

Vidal, J. P. (1998). The effect of emigration on human capital formation. *Journal of Population Economics*, *11*(4), 589-600. https://doi.org/10.1007/s001480050086

Vogler, M., & Rotte, R. (2000). The effects of development on migration: Theoretical issues and new empirical evidence. *Journal of Population Economics*, *13*(3), 485-508. https://doi.org/10.1007/s001480050148

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, *21*(6), 459-468. https://doi.org/10.1093/heapol/czl029

White, M. J., & Lindstrom, D. P. (2005). Internal Migration. In D. L. Poston & M. Micklin (Eds.), *Handbook of Population* (pp. 311-346). Springer US. https://doi.org/10.1007/0-387-23106-4_12

Yankow, J. J. (1999). The Wage Dynamics of Internal Migration within the United States. *Eastern Economic Journal*, *25*(3), 265-278.

ORCID

Rodrigo Pérez Silva https://orcid.org/0000-0003-3100-1658

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 129-148
https://doi.org/10.38191/iirr-jorr.24.062

Articles

Job quality in the shadow of informality: the mediating role of education. A regional analysis in Ecuador between 2014 to 2019

Viviana Carriel, César Andrés Mendoza**, Rodrigo Mendieta***, Sofía Bravo*****

Received: 20 February 2024
Accepted: 19 September 2024

Abstract:

This paper investigates the relationship between informality and job quality in Ecuador, emphasizing the role of education as a mediating factor. Utilizing data from the National Survey of Employment, Unemployment, and Underemployment from 2014 to 2019, the study employs an Employment Quality Index to assess job quality. Through a detailed empirical strategy incorporating pooled ordinary least squares regressions, panel data analyses, and spatial models, the research unveils a negative impact of informality in employment quality. However, the findings indicate that education does not significantly influence this relationship at the regional level. These results highlight the necessity for targeted policy interventions that address the structural issues in the labor market.

Keywords: Employment quality index; informality; education; Ecuador.
JEL Classification: J46; R23.

Calidad del empleo en la sombra de la informalidad: un análisis regional en Ecuador entre 2014 y 2019

Resumen:

Este artículo examina la relación entre la informalidad y la calidad del empleo en Ecuador, destacando el papel de la educación como factor mediador. Utilizando datos de la Encuesta Nacional de Empleo, Desempleo y Subempleo de 2014 a 2020, el estudio emplea un Índice de Calidad de Empleo para evaluar la calidad del empleo. A través de una estrategia empírica detallada que incorpora regresiones de mínimos cuadrados ordinarios agrupados, análisis de datos de panel y modelos de rezago espacial, la investigación revela un impacto negativo de la informalidad en la calidad del empleo. Sin embargo, los resultados indican que la educación no influye en esta relación a nivel regional. Los hallazgos subrayan la necesidad de intervenciones de política específicas, que aborden los problemas estructurales del mercado laboral.

Palabras clave: Índice de Calidad de empleo; informalidad; educación; Ecuador.
Clasificación JEL: J46; R23.

* Facultad de Ciencias Económicas y Administrativas, Universidad de Las Américas, Quito. Ecuador. / GIER – Facultad de Ciencias Económicas y Administrativas, Universidad de Cuenca, Cuenca. Ecuador. viviana.carriel@udla.edu.ec
** Instituto de Altos Estudios Nacionales – Escuela de Economía Pública y Sectores Estratégicos, Quito. Ecuador. / GIER – Facultad de Ciencias Económicas y Administrativas, Universidad de Cuenca, Cuenca. Ecuador. cesar.mendoza@ucuenca.edu.ec
*** GIER – Facultad de Ciencias Económicas y Administrativas, Universidad de Cuenca, Cuenca. Ecuador. rodrigo.mendieta@ucuenca.edu.ec
**** Facultad de Ciencias Económicas y Administrativas, Universidad de Cuenca, Cuenca. Ecuador. sofia.bravo@ucuenca.edu.ec
Corresponding author: viviana.carriel@udla.edu.ec

1. Introduction

In developing countries, informality in labor markets has been widely studied. Among the most notable studies in this area are those that analyze the relationship between informality and development (La Porta and Shleifer, 2014; Loayza, 2016; Ulyssea, 2018; Ulyssea, 2020); the characteristics of informal workers (Portes et al., 1986; Maloney, 2004; Bernal, 2009); the effects on the productivity of companies (Chong et al., 2008; Benjamin and Mbaye, 2012; Ulyssea, 2018); and its social consequences (Saavedra and Tommasi, 2007; Kanbur, 2017).

Another area of study is the relationship between informality and the quality of employment. Studies show that informality decreases the quality of employment (Mahmood, 2018; Neog and Sahoo, 2021; Apablaza et al., 2023). However, there are nuances within this relationship. Given that for certain workers informality may be voluntary, this does not correspond in all cases to low quality of employment (Pagés and Madrigal, 2008). Additionally, there are other factors that can mediate the relationship between informality and employment quality, such as education. Workers with more education can benefit from better quality jobs (Conover et al., 2022).

The aim of this article is to analyze the regional effect of informality on employment quality and how regional educational attainment moderates this relationship, an area that has been little studied in the literature. This study is particularly valuable as it focuses on Ecuador, a country that has not yet been thoroughly analyzed in the context of informality and job quality. The lack of existing research on Ecuador presents a unique opportunity to explore these dynamics in a region where labor market conditions and economic structures differ significantly from other Latin American countries. By shedding light on Ecuador's specific context, this research contributes new insights to the broader discussion on employment quality and informality.

We chose to use aggregated regional data from the ENEMDU database to capture broader interactions and neighborhood effects to understand how higher educational attainment moderates the informality impact on overall job quality. The Employment Quality index (EQI) is constructed following the methodology proposed by Apablaza et al. (2023), which in turn uses the Alkire Foster method (Sehnbruch et al., 2020) to work on the different dimensions of worker deprivation. Taking the EQI as the dependent variable and among the explanatory variables the informality and education of the population, several econometric models are used to estimate the effect of informality on the EQI in the regions of Ecuador mediated through education, including spatial models.

Workers with higher education are generally expected to moderate the negative effects of informality on job quality because their skills and qualifications should lead to better employment opportunities. In the Ecuadorian context, a negative effect of the tertiary education rate on the informality rate is evident (Appendix 1), in line with works such as that of Ulyssea (2020). However, our empirical evidence from the Ecuadorian context suggests a more intricate scenario where higher education levels do not mitigate the negative effect of informality on job quality.

This can be attributed to a labor market mismatch, where highly skilled workers end up in lower-skilled positions within the informal sector due to a scarcity of high-skill job opportunities (Battu and Bender, 2020). In fact, Leuven and Oosterbeek (2011) estimate that the proportion of overeducated people in Latin America is higher than that of undereducated people. In Ecuador, overeducation was close to 30% in 2017, with higher rates in agricultural and elementary occupations (Castro et al., 2024).

Although the evidence on the relationship between informality and overeducation is limited (Battu and Bender, 2020), overeducation may positively correlate with informality as efforts to elevate educational levels in the population have not been matched by the labor market's ability to generate suitable jobs. Particularly in Ecuador, where the presence of multinational corporations is limited, and foreign direct investment (FDI) is low (Camacho and Bajaña, 2020).

Moreover, informal practices are deeply entrenched in the labor markets of many economies (Jones et al., 2006; Contreras et al., 2008), which may limit the potential of education to mediate the relationship between informality and job quality. Indeed, evidence from Peru suggests that the problem in reducing

the informal sector is not the lack of supply of better educated workers, but the shortage of educated entrepreneurs capable of managing productive businesses (La Porta and Shleifer, 2014).

Moving forward, the paper is structured as follows: section 2 synthesizes existing research on the dynamics of informality and job quality, section 3 outlines the methodological approach adopted for constructing the EQI and delineates the regression models used to analyze the impact of informality, section 4 provides an overview of the data source, section 5 presents the empirical findings, offering a comparative analysis of informality rates and EQI across different time periods and regions of Ecuador, supported by spatial analysis techniques, section 6 discusses the outcomes of the regression analyses and section 7 draws together the key insights from the study and implications for policy.

2. Literature review

The dynamics of the labor market, particularly in relation to informality, have long been a central focus of economic research, especially within the context of developing countries. A critical aspect emerging from this research is the relationship between informality and the quality of employment. Ulyssea (2020) points out that transitions into and out of informality often involve young, low-skilled workers and women, suggesting a pattern where vulnerable groups are more likely to engage in informal employment. This observation underscores the need to understand the pathways through which individuals enter and exit informal work, and how these transitions affect job quality.

Neog and Sahoo (2021) illustrates a high concurrence of informality rates with poor-quality jobs in India, particularly in the unorganized sector, highlighting the pervasiveness of informality and poor job quality. This overlap suggests that various criteria used to define informality consistently identify a similar group of workers who are often engaged in low-quality employment.

Among the studies that have been developed for developing countries is that of Mahmood (2018), in which a strong relationship is identified between the shift from informal to formal employment and improvements in job quality. This transition is often linked to increased productivity and higher per capita incomes, suggesting that structural changes in employment can lead to a significant reduction in vulnerability and an enhancement in the quality of jobs.

Pagés and Madrigal (2008) offer a nuanced perspective, suggesting that for some workers, informality might be a choice rather than a consequence of exclusion. Their analysis of job satisfaction data from three low-income countries reveals that job satisfaction varies significantly within different types of informal employment, indicating that informality does not uniformly equate to lower job quality across all contexts. The global review by Sethuraman (1998) emphasizes the varying quality of employment within the informal sector. They point out that while informal employment is generally of inferior quality compared to formal employment, there is considerable variation within the informal sector itself.

In the context of Latin American countries, Bertranou et al. (2013) explore the segmentation of the labor market in Argentina, highlighting the prevalence of precarious employment in various forms, including informal salaried employment and unskilled self-employment. Their analysis highlight that the most widespread factor affecting job quality in Argentina is informality, shaped by a complex interplay of economic, institutional, and regulatory factors.

The relationship between informality and job quality is further illuminated by the work of Apablaza et al. (2023), which provides a comprehensive analysis of this dynamic in Chile. This study utilizes a deprivation score to reflect the breadth of deprivations across various dimensions of job quality, offering a nuanced understanding of the factors contributing to poor quality of employment. One of the key findings of the study is that the most significant improvements in quality of employment deprivation across all regions are related to changes in working hours and income indicators. These improvements were largely driven by legal changes, particularly in working hours and minimum wage regulations, highlighting the profound impact of policy on employment quality. However, it is interesting to note that there are other factors that influence the quality of employment. Thus, regions with larger expansions in education showed stronger improvements in job quality, indicating a direct correlation between educational attainment and employment quality.

In this line of research, Chong et al. (2008) discovered in Peru that the training program for young workers (PROJOVEN) positively impacts the probability of securing high-quality employment. It is important to highlight that Peru, along with other Latin American countries like Ecuador, contends with labor markets characterized by high levels of informality and spatial inequality (Guevara-Rosero and Del Pozo, 2020). In Mexico, Conover et al. (2022) find evidence that workers with higher levels of education are more likely to start in the formal sector and remain in it, as well as to transition from the informal sector to the formal sector, benefiting from better salaries and higher job quality.

These studies support the hypothesis that education helps workers obtain higher quality jobs that are related to the formal sector, despite being in labor markets with high rates of informality. However, there is evidence that highly skilled workers end up in lower quality informal jobs, for which they are overeducated. Handel et al. (2016, 2017) find for low- and middle-income countries a positive relationship between informality and overeducation, arguing that this is due to the lack of job options.

This research further delves into this complex relationship, emphasizing education's role not just as a mediator but also as a potential exacerbator of the negative impacts of informality on job quality. While higher educational attainment is traditionally seen as a pathway to better employment opportunities, this study suggests a paradox where increased education levels may also lead to greater underemployment in informal sectors. This phenomenon points to a mismatch between the labor market's evolution and educational advancements, particularly in contexts like Ecuador, where the labor market is marked by high informality (Mendoza and Jara, 2020), limited multinational corporations, and low foreign direct investment (Camacho and Bajaña, 2020).

In line with this, La Porta and Shleifer (2014) indicate that in fact to reduce the informal sector and its practices it is necessary to have a supply of educated entrepreneurs and not only educated workers. This could be one of the factors why the level of education of workers does not counteract the effects of informality on the quality of employment. This scenario highlights the need for targeted policy interventions that not only enhance educational outcomes but also address the scarcity of educated entrepreneurs capable of creating competitive businesses. Policies should also aim to align educational achievements with labor market demands, particularly in regions where informal practices are deeply entrenched, to mitigate the adverse effects of informality on job quality."

3. Empirical strategy

3.1. Employment quality index

In OECD countries, the quality of employment is measured by three dimensions: income, security in the labor market and quality of the work environment; for which specific surveys that address these dimensions are required (Cazes et al., 2015). In the case of Ecuador, it is very difficult to obtain information on the quality of the work environment with the available surveys. Therefore, our approach follows the methodology established by Apablaza et al. (2023). This methodology, which adeptly captures the multifaceted nature of employment quality, is particularly well-suited for examining the Ecuadorian labor market in its diverse and nuanced forms.

In constructing an Employment Quality Index (EQI) for Ecuador, we adopt the Alkire-Foster method (Sehnbruch et al., 2020), a robust and versatile tool for measuring multidimensional deprivations. This method, known for its application in assessing poverty, lends itself remarkably well to the study of employment quality. It allows us to count multiple types of deprivations experienced by individuals simultaneously and to weigh these deprivations equally or differently depending on their relative importance. The flexibility of this method is key; it can be adapted to various contexts and thresholds, making it suitable for tracking changes over time and across different groups.

Following Apablaza et al. (2023), our application of the Alkire-Foster method revolves around three core dimensions of employment quality: labor income, employment stability, and working conditions. It should be noted that these dimensions have a certain relationship with those used in OECD countries and

in the case of the dimension of quality of the work environment, this is only measured based on the number of hours worked per week, which in the OECD framework is captured together with other indicators.

Each dimension is accorded equal weight in the overall index, with each comprising a set of specific indicators (Table 1). Equal weights were assigned to ensure a balanced representation of the various dimensions of job quality. This approach facilitates meaningful comparisons across regions. However, while equal weighting avoids bias toward any particular dimension, it may also oversimplify the complex trade-offs that workers make between different aspects of employment, such as income versus job stability. This limitation, though common in multidimensional indices, highlights the potential value of exploring alternative weighting schemes in future research. Nonetheless, equal weighting remains consistent with established practices in the literature and serves as a foundational framework for understanding the multifaceted nature of job quality.

Labor income, for instance, is gauged through hourly income, benchmarked against the basic hourly wage. The rationale behind benchmarking this indicator against the basic hourly wage is deeply rooted in the national labor regulations of Ecuador, the State annually establishes the unified basic wage for private sector workers (Ministerio de Trabajo, 2022). This benchmark reflects a nationally recognized standard and ensures that our assessment of labor income in terms of employment quality is grounded in the economic realities and legal standards prevailing in the country.

Employment stability encompasses indicators such as occupational status and job tenure, distinguishing between contract workers and self-employed workers (Apablaza et al., 2023), and accounting for the duration of employment in years. The job tenure indicator uses a threshold of one year, reflecting significant milestones in an Ecuadorian worker's journey. After completing a year of service, employees not only become eligible for annual vacations but also receive their first payment of the thirteenth and fourteenth salaries, marking important financial benefits. Moreover, aligning with Article 196 of the Ecuadorian Labor Code this one-year milestone entitles workers to an accumulation in their reserve fund or capitalized labor, equivalent to a month's salary for each subsequent year of service (Registro Oficial, 2020).

Working conditions are assessed through indicators like social security contributions and working hours, set against thresholds for social security coverage and a standard 40-hour workweek, legally in force since 1997 (Zamora, 2013).

Table 1.
Dimensions and Indicators of the Employment Quality Index (EQI)

Dimension	Indicator	Weight	Thresholds
Labor Income	Hourly Income	1/3	Less than basic hourly wage
Employment Stability	Occupational Status	1/6	No contract, self employed
	Job Tenure (years worked)	1/6	Less than 1 years
Working Conditions	Social Security	1/6	No affiliation to the pension system
	Working Hours	1/6	More than 40 hours per week

Source: Own elaboration

The Alkire-Foster method is then employed to calculate the index. This method first calculates the proportion of individuals who are deprived in multiple dimensions (incidence) and then assesses the average deprivation score among these individuals (intensity). This dual focus provides a more nuanced understanding of employment quality, as it captures not only how many workers are experiencing poor job quality but also the depth of their deprivation. It should be noted that this study differs from that of Orellana et al. (2020) as they use polychoric principal component analysis to construct a job quality index.

In the Alkire-Foster framework, let's consider a population N of individuals, each potentially experiencing deprivations in various dimensions of employment quality. We define D as the set of dimensions. For each individual i in the population, and each dimension d in D, we identify whether the individual is deprived in that dimension. If individual i is deprived in dimension d, we denote this as $d_i = 1$; otherwise, $d_i = 0$. Each dimension d is assigned a weight w_d.

For each individual, we calculate a deprivation score, c_i, which is the sum of the weights of the dimensions in which the individual is deprived: $c_i = \sum_{d \in D} d_i * w_d$. This score represents the breadth of an individual's deprivations across all considered dimensions. A cutoff k is set, which is a threshold that determines when an individual is considered to have poor quality employment. This could be, for instance, being deprived in at least one of the key dimensions.

Finally, two key indices are calculated: i) the headcount ratio (H), and ii) the average Intensity of deprivation (A). H is the proportion of the population that is deprived in their quality of employment (i.e., those for whom $c_i \geq k$). A is the average deprivation score among those who are considered to have poor quality employment.

The product between H and A (known as adjusted headcount ratio (M_0)) provides a comprehensive measure that reflects both the prevalence and the intensity of poor employment quality in the population. Finally, to calculate the EQI, the inverse of M_0 is considered, in this way values close to 1 represent higher job quality.

3.2. Informality and employment quality

Our primary objective is to estimate the regional impact of informality employment, understood as workers in companies or businesses that are not registered with the tax authority, and therefore do not pay taxes (Mendoza and Jara, 2020) on the quality of employment, as measured by the Employment Quality Index (EQI). To achieve this, we will employ a regression model that accounts for the panel structure of the data. The EQI serves as the dependent variable, aggregated at the regional level. Informality, measured as the proportion of the informal workforce in each region, serves as the primary independent variable. The panel data regression equation can be expressed as:

$$EQI_{it} = \alpha + \beta * Informality_{it} + \gamma X_{it} + \mu_i + \lambda_t + \epsilon_{it}$$

Where EQI_{it} is the Employment Quality Index for region i at time t. $Informality_{it}$ is the informality rate, X_{it} is a vector of control variables. α is the intercept, β is the coefficient measuring the impact of informality on EQI, γ is a vector of coefficients for the control variables, μ_i represents region-specific fixed effects to control for unobservable characteristics that are constant over time for each region but vary across regions, λ_t represents time-specific fixed effects to control for factors that vary over time but are constant across all regions, such as national economic policies or global economic trends, ϵ_{it} is the error term.

Incorporating insights from previous studies (Apablaza et al., 2023; Bertranou et al., 2013), the model includes several control variables at the regional level. Variables such as regional GDP per capita, educational and sectoral composition of the workforce in each region are considered. To thoroughly assess the influence of regional educational attainment on the relationship between informality and employment quality, we incorporate another specification that considers an interaction term between informality and education. This approach is particularly crucial in understanding how the regional disparities in education levels across Ecuador can impact the way in which informality affects employment quality.

This interaction term is designed to capture the differential effects of informality on employment quality based on the varying levels of educational attainment across regions. The coefficient of this interaction term is key to our analysis. It indicates the extent to which the impact of informality on employment quality is moderated by education. A significant value of this coefficient in our model would suggest that the level of education in a region can either amplify or mitigate the effects of informality on job quality. For instance, regions with higher educational attainment might exhibit a weaker negative relationship between informality and employment quality. This observation is supported by findings from

Conover et al. (2022), which indicates that better-educated workers tend to have a higher likelihood of securing formal and better-quality jobs, and also they are more inclined to start in formal employment, persist in these roles, and transition to them from informal jobs, also benefiting more significantly in terms of wages from such transitions.

This in-depth analysis is not just a statistical exercise but has profound policy implications. If educational attainment is found to significantly mediate the impact of informality, this would underscore the importance of educational policies in mitigating the negative aspects of informality. It could indicate that regions with higher education levels are better equipped to transform informal employment into more quality-driven job opportunities, or conversely, that regions with lower education levels might be more vulnerable to the negative effects of informality.

In the other hand, if higher education exacerbates informality's adverse effects, it signals a misalignment between the labor market's needs and the educational system's output. This knowledge is vital for designing educational policies that not only improve educational attainment but also ensure that the skills acquired are relevant and valued in the labor market.

4. Data description

The backbone of our empirical analysis is the data derived from the National Survey of Employment, Unemployment, and Underemployment (ENEMDU), conducted by Ecuador's National Institute of Statistics and Census (INEC) from 2014 to 2019. This dataset offers a comprehensive and granular view of the Ecuadorian workforce, encompassing a wide range of variables such as gender, age, household head status, social security coverage, marital status, ethnicity, educational level, working hours, years of employment, and occupational status. The richness of ENEMDU data, coupled with its longitudinal nature, provides an unprecedented opportunity to delve into the nuances of employment quality in Ecuador.

In our approach, we have strategically chosen to use the ENEMDU surveys from June and December of each year, as these tend to have a larger sample size, offering greater representativeness at the regional level. This decision enhances the robustness of our analysis, ensuring that the insights we derive are reflective of the diverse conditions across the various regions of Ecuador. Additionally, the starting point of our dataset, 2014, is significant as it marks the year when the surveys began to offer distinct representation for the regions of the Amazonia (MIES, 2017). Prior to this, the representation of these regions was aggregated, limiting the granularity of regional analysis.

It is important to note that all variables used in this analysis are sourced from the National Survey of Employment, Unemployment, and Underemployment (ENEMDU), except for the Gross Domestic Product (GDP), which is obtained from the Central Bank of Ecuador. Thus, following Apablaza et al. (2023) and Bertranou et al. (2013), some variables are incorporated into the model. The informality rate is measured as the ratio between the number of workers in informal companies or enterprises and the total number of workers in the regions. It should be noted that since 2015, the new methodology used to calculate informal employment considers companies with less than 100 workers that are not registered in the Single Registry of Taxpayers as informal; therefore, workers in these companies are considered informal (Vera et al., 2015). With this, an update and homologation were carried out in the ENEMDU databases, which provide a measure of informality since 2007 with the new methodology.

The gross employment rate is measured as the ratio of employed to the economically active population. The population with tertiary education is calculated as a dichotomous variable, which takes values of 1 when the rate of population with tertiary education is higher than the national rate. The rate of workers in microenterprises is measured as the ratio of the number of workers in microenterprises (enterprises with less than 10 workers) to the total number of workers[1]. The rate of workers in the tertiary

[1] The definition of microenterprises is based on two dimensions: the number of workers and the sales volume. Thus, companies with less than 10 workers and sales that do not exceed $100,000 per year are considered microenterprises (INEC, 2021). However, the ENEMDU does not have information on the sales volume of companies, which is why authors such as Guevara-Rosero (2021) only work with the criterion of the number of employees.

sector is calculated as the proportion of workers in the service sectors (branches from G to U of the International Standard Industrial Classification). The urban population rate is calculated as the proportion of the population residing in urban areas. Finally, GDP per capita is measured by dividing GDP by the total population in each region, assuming that the regions' share of total GDP remains the same between June and December of each year.

Additionally, meticulous processing of the ENEMDU data is carried out to align it with the dimensions and indicators of the EQI. The analysis is finely aggregated to the provincial level, allowing us to uncover regional variations in employment quality across Ecuador. Subsequently, control variables that influence the quality of employment are taken into account. Table 2 shows a detailed description and descriptive statistics of these variables.

Since 2012, Ecuador has comprised 24 provinces. Considering the analysis period from 2014 to 2019 (12 periods), the dataset includes a total of 288 observations However, in June 2016 and 2017, and in December 2018, the databases do not contain information for the province of Galapagos, so the final database size is 285 observations. Thus, the panel dataset used in this study consists of 24 cross-sectional units corresponding to Ecuador's provinces, observed over 12 time periods. However, due to missing data for the Galápagos province in certain years, the dataset is classified as unbalanced.

To avoid multicollinearity, the correlation between these variables was calculated, showing that the rate of people employed in the tertiary sector and the rate of people living in urban areas showed the highest correlation (0.86). Additionally, the variance inflation factor was calculated where the rate of people employed in the tertiary sector showed the highest value (Appendix 2). With these considerations, this variable was excluded from the analysis.

TABLE 2.
Descriptive statistics of aggregated variables used in the estimates

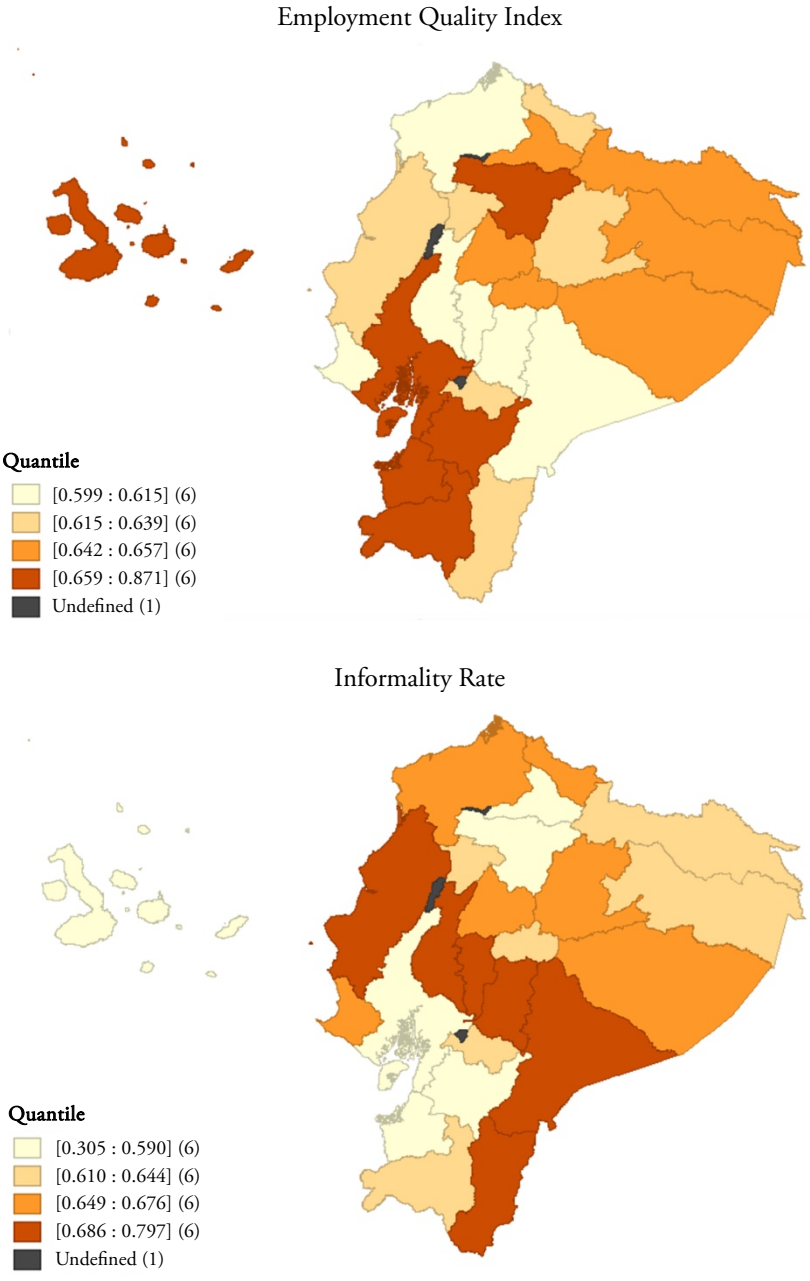
Variable	Definition	Min	Max	Mean	Std. Dev	Obs
EQI	$1 - M_0$	0.5450	0.8990	0.6484	0.0581	285
Rate of informality	$\frac{employees\ in\ informal\ sector_{i,t}}{employees_{i,t}}$	0.1277	0.8580	0.4862	0.1254	285
GDP per capita	$\frac{GDP_{i,t}}{population_{i,t}}$	1.0555	27.7145	2.8356	2.8296	285
Gross employment rate	$\frac{employees_{i,t}}{economically\ active\ population_{i,t}}$	0.5053	0.8758	0.6674	0.0785	285
Population with tertiary education	1 if the rate of population with tertiary education is higher than national average	0	1	0.4175	0.4940	285
Rate of employees in microenterprises	$\frac{employees\ in\ microenterprises_{i,t}}{employees_{i,t}}$	0.7498	0.9899	0.9021	0.0546	285
Rate of employees in the tertiary sector	$\frac{temployees\ in\ the\ tertiary\ sector_{i,t}}{employees_{i,t}}$	0.1430	0.8248	0.4640	0.1246	285
Urban population rate	$\frac{urban\ population\ _{i,t}}{population_{i,t}}$	0.1207	0.9652	0.5379	0.2113	285

Source: ENEMD

5. Informality and employment quality index in Ecuador

The comparative analysis of informality and job quality in Ecuador for 2014 and 2019 reveals profound regional disparities during the period. Provinces such as Pichincha, Galapagos, Guayas, Azuay and El Oro show lower informality rates and higher employment quality indices, potentially indicating more robust formal employment structures and better job conditions (Figure 1). Conversely, provinces like Bolívar, Los Ríos, Chimborazo and Morona Santiago exhibit higher informality rates, suggesting a prevalence of informal employment and potentially lower job quality.

Figure 1.
Informality Rate and Employment Quality in Ecuadorian Provinces



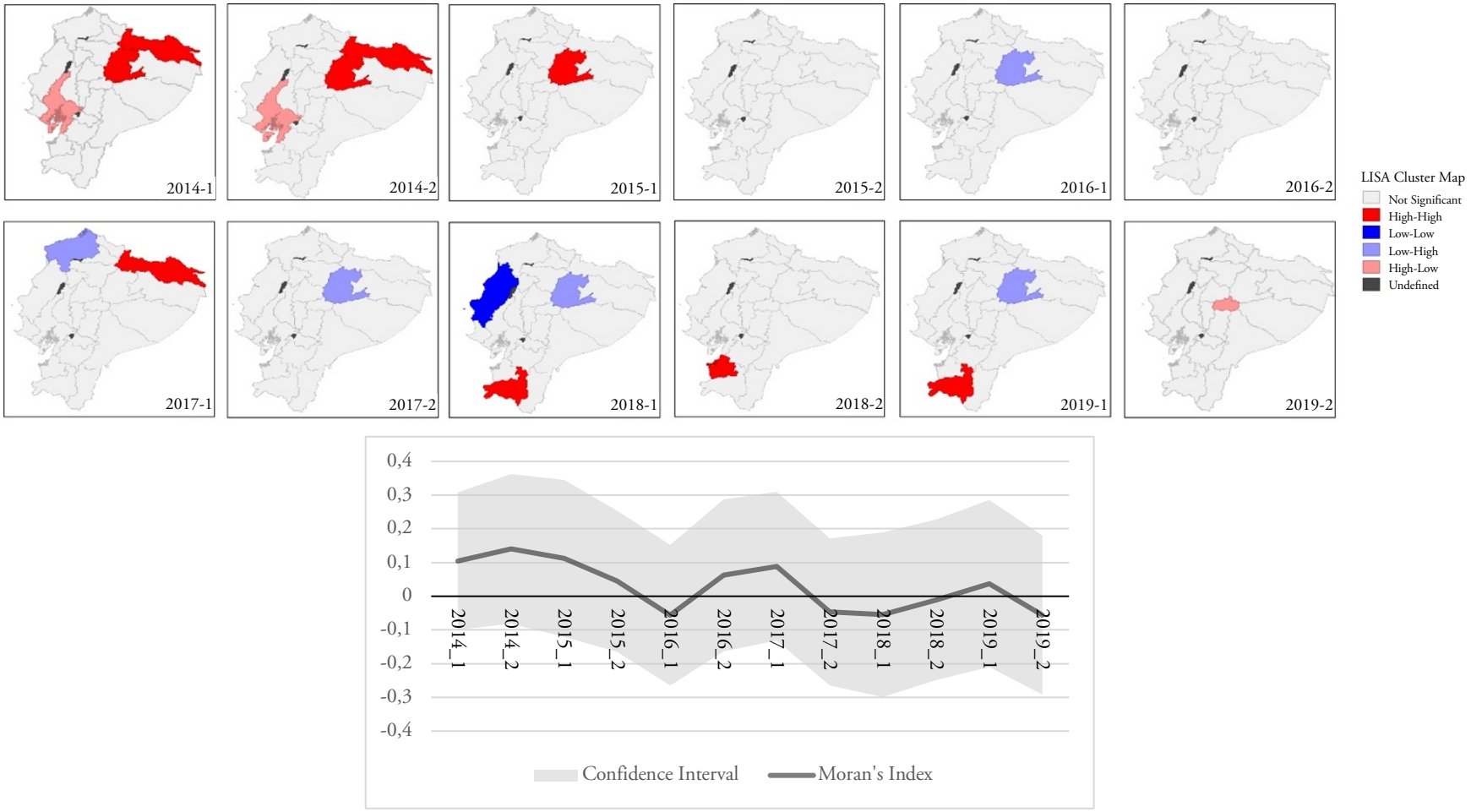
Note: Average for the period 2014 to 2019.
Source: ENEMDU.

The Local Indicators of Spatial Association (LISA) cluster maps and Moran's I statistics for the Employment Quality Index (EQI) provide a localized and global measure of spatial autocorrelation, respectively, for the years 2014 and 2019 (Figure 2). The LISA cluster map shows between 2014 and 2015 clusters with provinces in the north of the country where the EQI takes high values that are surrounded by provinces with similar values (red provinces on the map). During 2016 this pattern disappears and between 2017 and 2018 it is observed that in this area of the country there are provinces with low EQI values surrounded by provinces with equally low values (the case of Esmeraldas in the far north and Sucumbíos towards the northeast) or provinces with high EQI values (the case of Manabí in the northwest of the country). Finally, in 2018 and 2019, clusters of provinces with high EQI values surrounded by provinces with equal values appear again, but now in the southern part of the country. Despite these localized patterns in certain areas of the country, the Moran's I shows that there is no spatial autocorrelation for the EQI during the period analyzed. That is, it can be assumed that the distribution of EQI values is random between provinces.

The analysis of the spatial distribution of informality rates in Ecuador through LISA cluster maps and Moran's I statistics from 2014 to 2019 (Figure 3) shows clusters of provinces where the informality rate is low, in the northern part of the country. This contrasts with the fact that in these areas there are also clusters of provinces with high EQI values. However, this pattern fades in 2016, 2018 and 2019. Additionally, the Moran index shows that there is no spatial autocorrelation for this indicator, except for the second half of 2014 where a positive spatial correlation is shown.

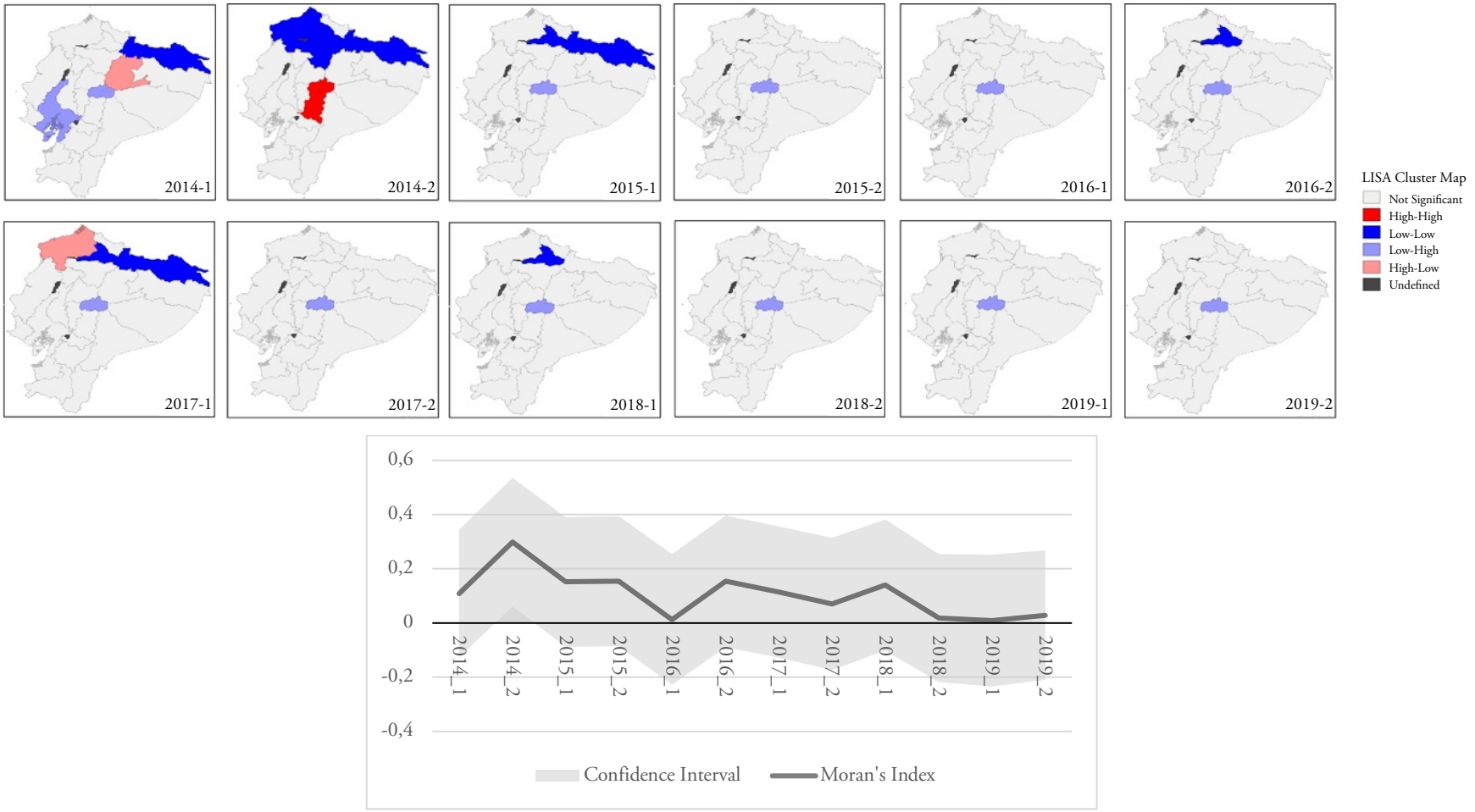
In the pursuit of understanding the broader interplay between informal employment and employment quality within Ecuadorian provinces, a bivariate LISA analysis was conducted. This analysis aimed to discern whether the informality rates of neighboring provinces exert an influence on the Employment Quality Index (EQI) within a given province. The results revealed that the relationship between a province's EQI and its neighbors' informality rates was not statistically significant. This suggests that the quality of employment within a province may not be directly affected by the informal employment rates in the adjoining regions. This finding indicates a level of independence in the factors determining the quality of employment, implying that local provincial conditions, policies, and economic dynamics may play a more substantial role than the influence of neighboring provinces' employment structures.

FIGURE 2.
LISA cluster map and Moran's I for Employment Quality Index



Source: ENEMDU.

FIGURE 3.
LISA cluster map and Moran's I for Informality



Source: ENEMDU

6. Results

We present an examination of the determinants of the Employment Quality Index (EQI) across Ecuadorian provinces, utilizing pooled ordinary least squares (OLS) regressions and panel data techniques. Our empirical strategy progressively incorporates additional complexities to elucidate the nuanced effects of informality on employment quality, especially considering the interaction with tertiary education levels relative to the national context.

Pooled OLS regression (Model 1) in Table 3, reveals a statistically significant negative relationship between the rate of informality and EQI. This negative coefficient suggests that an elevation in informality is associated with a detraction in employment quality, holding other factors constant.

Enhancing the robustness of our findings, we introduced fixed effects for both year and region in Model 2, adjusting for unobserved heterogeneity that could confound our estimates. In this specification, the coefficient for informality marginally attenuates to -0.2051, yet remains negatively significant, thereby affirming the adverse impact of informality on employment quality within fixed regional contexts.

In Model 3, we enriched our analysis by including an interaction term between informality rates and a binary indicator of tertiary education, which is set to 1 for provinces with a proportion of tertiary-educated population exceeding the national average. The interaction term emerged as significantly negative. This outcome indicates that the detrimental effect of informality on employment quality is more pronounced in provinces with higher levels of tertiary education.

Since the data we are working with follow a panel data specification, the Breusch Pagan test is applied to determine whether the random effects estimator with panel data is better than using the OLS estimator in pooled data. Then, the Hausman test is applied to decide the best panel data model between fixed and random effects (Appendix 3). After these tests, it is determined that the best model is the fixed effects model, which is shown in specification 4 of Table 3. In this model it is observed that the informality rate maintains its negative effect on the quality of employment, however the education and the interaction between education and informality are not significant.

Since there is some evidence of spatial autocorrelation in the informality rate and of localized spatial autocorrelation in both the employment quality index and the informality rate, it was decided to additionally estimate the spatial error and spatial lag model (SAC model) with fixed effects to panel data. Across all models in Table 4, the rate of informality consistently exerts a negative impact on EQI, underscoring the detrimental effects of informal employment on job quality, even when accounting for spatial dependencies. In fact, the significance of both the lambda (λ) and the rho (ρ) in the SAC model suggests that spatial dependencies are crucial in understanding the relationship between informality and job quality in Ecuador. These findings indicate that not only do local factors within a region influence employment quality, but there are also significant spillover effects from neighboring regions.

However, even considering these relationships of the regions in space, it is evident that education is not influencing the relationship between informality and the quality of employment. These results could be due to structural factors, such as the lack of formal employment opportunities and the widespread informal practices that dominate the labor market, leading to a scenario in which education does not translate into better quality of employment. As a result, even highly educated individuals might not achieve better employment outcomes due to the entrenched nature of informal work.

TABLE 3.
OLS and Random Effects Panel Data Estimation Results

Dependent variable: EQI_{it}	Pooled OLS (1)		Pooled OLS with time-region effects (2)		Pooled OLS with time-region effects and interaction (3)		Panel fixed effects with interaction (4)	
	Coefficient (*Std. Dev.*)		Coefficient (*Std. Dev.*)		Coefficient (*Std. Dev.*)		Coefficient (*Std. Dev.*)	
Rate of informality	-0.4309 (0.0292)	***	-0.2051 (0.0328)	***	-0.1338 (0.0322)	***	-0.1485 (0.0385)	***
Population with tertiary education	0.0160 (0.0049)	***	0.0241 (0.0043)	***	0.1349 (0.0170)	***	0.0103 (0.0237)	
Interaction between informality and tertiary education					-0.2371 (0.0354)	***	-0.0124 (0.0488)	
GDP per capita	0.0008 (0.0007)		0.0010 (0.0006)		0.0012 (0.0006)	**	-0.0002 (0.0009)	
Gross employment rate	0.1873 (0.0317)	***	0.0166 (0.0349)		-0.0245 (0.0329)		-0.1255 (0.0400)	***
Rate of employees in microenterprises	0.0197 (0.0655)		-0.1106 (0.0644)	*	0.0044 (0.0621)		-0.2096 (0.0671)	***
Urban population rate	-0.0647 (0.0157)	***	-0.0265 (0.0152)	*	-0.0237 (0.0141)	*	0.0197 (0.0243)	
Constant	0.7411 (0.0596)	***	0.8440 (0.0590)	***	0.7243 (0.0575)	***	0.9776 (0.0632)	***
Time fixed effects			*Yes*		*Yes*		*Yes*	
Region fixed effects			*Yes*		*Yes*		-	
# observations	285		285		285		285	
Adj. R squared	0.7035		0.7985		0.8272		-	
Root MSE	0.0317		0.0261		0.0242		-	

Source: Own elaboration.

TABLE 4.
SAC Model Estimation Results

Dependent variable: EQI_{it}	SAC Model with fixed effects (1)		SAC Model with fixed effects and interaction (2)	
	Coefficient (*Std. Dev.*)		Coefficient (*Std. Dev.*)	
Rate of informality	-0.1140 (0.0342)	***	-0.1141 (0.0350)	***
Population with tertiary education	0.0052 (0.0039)		0.0051 (0.0213)	
Interaction between informality and tertiary education			0.0002 (0.0441)	
GDP per capita	-0.0004 (0.0008)		-0.0004 (0.0008)	
Gross employment rate	-0.1517 (0.0356)	***	-0.1517 (0.0358)	***
Rate of employees in microenterprises	-0.2271 (0.0605)	***	-0.2271 (0.0606)	***
Urban population rate	0.0323 (0.0232)		0.0323 (0.0233)	
Time Fixed Effects	Yes		Yes	
# observations	276		276	
Overall R squared	0.1848		0.1847	
Rho	0.3526 (0.1706)	**	0.3528 (0.1757)	**
Lambda	-0.4142 (0.2393)	*	-0.4146 (0.2489)	*

Source: Own elaboration

7. CONCLUSIONS

In light of the findings from this research on the Ecuadorian labor market, we understand that the presence of informality significantly impacts the Employment Quality Index (EQI). Contrary to the expectation that higher education would mitigate the negative effects of informality on job quality, the study reveals that the presence of a more educated population does not appear to counterbalance the detrimental effects of informality on job quality.

A reason for this result could be the mismatch in the labors markets, if tertiary education is not aligned with industry needs, even a well-educated workforce may find themselves in informal sectors (Battu and Bender, 2020). Alternatively, the existence of deeply entrenched informal work practices might distort labor markets, suppressing wages and job quality even for those with advanced skills. Another reason for this result is that it would be educated entrepreneurs who can establish productive businesses against which informal businesses would not be able to compete, leading to a reduction in informality (La Porta and Shleifer, 2014) and thus improving the quality of work.

For policymakers, these findings highlight the critical need for strategies that align educational outcomes with labor market demands. This involves not just enhancing access to education but ensuring that education is directly linked to the needs of the economy. There is also a pressing need to foster entrepreneurship and support the development of businesses that can offer high-quality, formal employment opportunities.

Governments should address the issue of informal employment through a combination of strategies. While stricter enforcement of labor regulations and adjustments to minimum wage policies might be necessary, these should be complemented by efforts to expand formal job opportunities and improve the relevance of educational programs. The disconnect between education and employment outcomes in Ecuador and the fact that more education does not necessarily translate into more skills (Castro et al., 2024) indicates that reforms in the university system may be required, potentially focusing on practical skills, entrepreneurship, and stronger collaboration with industries to ensure graduates are prepared for the realities of the labor market.

The failure to tackle these issues could lead to significant consequences, including a 'brain drain' where highly educated individuals migrate to regions with better employment prospects, exacerbating regional inequalities. Furthermore, the persistence of underemployment could result in increased social dissatisfaction, as individuals find themselves in jobs that do not fully utilize their skills and education.

Recognizing the limitations of this study, particularly regarding data scope and potential confounding factors, future research should explore the causal pathways between education, informality, and job quality. Longitudinal studies, as well as comparative analyses across different regions or countries, could provide deeper insights into these complex dynamics, informing more effective policy interventions that can better address the challenges identified in this research.

References

Apablaza, M., Sehnbruch, K., González, P., & Méndez, R. (2023). Regional inequality in multidimensional quality of employment: insights from Chile, 1996–2017. *Regional Studies*, *57*(3), 416-433.

Battu, H., & Bender, K. A. (2020). Educational mismatch in developing countries: A review of the existing evidence. *The Economics of Education* (pp. 269-289). Academic Press.

Benjamin, N. C., & Mbaye, A. A. (2012). The Informal Sector, Productivity, and Enforcement in W est A frica: A Firm-level Analysis. *Review of Development Economics*, *16*(4), 664-680.

Bernal, S. (2009). The informal labor market in Colombia: Identification and characterization. *Desarrollo y sociedad*, (63), 145-208.

Bertranou, F., Casanova, L., Jiménez, M., & Jiménez, M. (2013). Informality and Employment Quality in Argentina: country case study on labor market segmentation. (Conditions of work and employment series, no. 49, ILO).

Camacho, F. R., & Bajaña, Y. S. (2020). Impact of foreign direct investment on economic growth: Comparative analysis in Ecuador, Peru and Colombia 1996–2016. *International Journal of Economics and Financial Issues*, *10*(4), 247-257.

Castro, J., Ortega, L., Yamada, G., & Mata, D. (2024). Overeducation and overskilling in Latin America: evidence from PIAAC. *Comparative Education*, *60*(2), 296-314.

Cazes, S., Hijzen, A., & Saint-Martin, A. (2015). *Measuring and assessing job quality: the OECD job quality framework.*

Chong, A., Galdo, J., & Saavedra, J. (2008). Informality and productivity in the labor market in Peru. *Journal of Economic Policy Reform*, *11*(4), 229-245.

Conover, E., Khamis, M., & Pearlman, S. (2022). Job quality and labour market transitions: Evidence from Mexican informal and formal workers. *The Journal of Development Studies*, *58*(7), 1332-1348.

Contreras, D., De Mello, L., & Puentes, E. (2008). *Tackling business and labour informality in Chile.* (OECD Working Paper).

Guevara-Rosero, G., & Del Pozo, D. (2020). Determination of the urban wage premium in Ecuador. *Investigaciones Regionales-Journal of Regional Research*, (47), 57-77.

Guevara-Rosero, G. C. (2021). Determinants of manufacturing micro firms' productivity in Ecuador. Do industry and canton where they operate matter? *Regional Science Policy & Practice*, *13*(4), 1215-1249.

Handel, M., Comyn, P., & Strietska-Ilina, O. (2017). Predictors and consequences of mismatch in developing countries: Results from the World Bank STEP survey. In *Skills and jobs mismatches in low and middle income countries* (pp. 63-87).

Handel, M., Valerio, A., & Puerta, M. L. S. (2016). *Accounting for mismatch in low-and middle-income countries: Measurement, magnitudes, and explanations.* World Bank Publications.

INEC. (2021). Directorio de Empresas y Establecimientos 2020. Available at: https://www.ecuadorencifras.gob.ec/documentos/web-inec/Estadisticas_Economicas/DirectorioEmpresas/Directorio_Empresas_2020/Principales_Resultados_DIEE_2020.pdf

Jones, T., Ram, M., & Edwards, P. (2006). Shades of grey in the informal economy. *International journal of sociology and social policy*, *26*(9/10), 357-373.

Kanbur, R. (2017). Informality: Causes, consequences and policy responses. *Review of Development Economics*, *21*(4), 939-961.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of economic perspectives*, *28*(3), 109-126.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market.. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 283-326). Elsevier Science.

Loayza, N. V. (2016). Informality in the Process of Development and Growth. *The World Economy*, *39*(12), 1856-1916.

Mahmood, M. (2018). Policy for Jobs: Reducing Informality. In *The Three Regularities in Development: Growth, Jobs and Macro Policy in Developing Countries* (pp. 233-270).

Maloney, W. F. (2004). Informality revisited. *World development*, *32*(7), 1159-1178.

Mendoza, C. A., & Jara, B. (2020). Natural Disasters and Informality: are local labor markets impacted after an earthquake? *Regional Science Policy & Practice*, *12*(1), 125-157.

Ministerio del Trabajo (2022). Acuerdo Ministerial No. MDT-2022-216. https://www.trabajo.gob.ec/wp-content/uploads/2024/01/ACUERDO-MINISTERIAL-Nro.-MDT-2022-216.pdf

MIES. (2017). Ficha metodológica. https://info.inclusion.gob.ec/index.php/fichas-metodologicas-peusrext?download=147:hogares-con-ninos-y-ninas-en-condicion-de-trabajo-infantil

Neog, B. J., & Sahoo, B. K. (2021). Defining and measuring informality in India. *Journal of Human Development and Capabilities*, *22*(3), 486-505.

Orellana, M., Rivera, C., Beltrán, P., & Ontaneda, D. (2020). Midiendo la calidad del empleo: una aplicación para Ecuador en el periodo de 2007 a 2017. *Revista de Economía del Caribe*, (25), 7-33.

Pagés, C., & Madrigal, L. (2008). Is Informality a good measure of job quality?: Evidence from job satisfaction data.

Portes, A., Blitzer, S., & Curtis, J. (1986). The urban informal sector in Uruguay: Its internal structure, characteristics, and effects. *World Development*, *14*(6), 727-741.

Registro Oficial. (2020). Código del Trabajo. https://www.ces.gob.ec/lotaip/2020/Junio/Literal_a2/C%C3%B3digo%20del%20Trabajo.pdf

Saavedra, J., & Tommasi, M. (2007). Informality, the State and the social contract in Latin America: A preliminary exploration. *International Labour Review*, *146*(3-4), 279-309.

Sehnbruch, K., González, P., Apablaza, M., Méndez, R., & Arriagada, V. (2020). The Quality of Employment (QoE) in nine Latin American countries: A multidimensional perspective. *World Development*, *127*, 104738.

Sethuraman, S. V. (1998). Gender, informality and poverty: A global review. Gender Bias in Female Informal Employment and Incomes in Developing Countries.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, *12*, 525-546.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, *108*(8), 2015-2047.

Vera, A., Álava, A., & Moncayo, J. (2015). Actualización metodológica: el empleo en el sector informal. *Revista de Estadística y Metodologías*, (1). 55-62.

Zamora Ulloa, J. I. (2013). Jornada parcial de trabajo como respuesta a la eliminación de la contratación laboral por horas en el Ecuador (Bachelor's thesis, Universidad del Azuay).

ORCID

Viviana Carriel https://orcid.org/0000-0001-6572-0648

César Andrés Mendoza https://orcid.org/0000-0003-3106-6742

Rodrigo Mendieta https://orcid.org/0000-0001-8040-9305

Appendixes

Appendix I

Dependent variable: Rate of informality$_{it}$	**Pooled OLS**	
	Coefficient (*Std. Dev.*)	
Rate of population with tertiary education	-0.3258 (0.0904)	***
Rate of employees in the tertiary sector	-0.4607 (0.0761)	***
Urban population rate	-0.0967 (0.0405)	**
Rate of employees in microenterprises	0.5856 (0.0947)	***
Gross employment rate	0.0926 (0.0573)	
GDP per capita	-0.0031 (0.0013)	**
Constant	0.1780 (0.0954)	*
Time fixed effects	Yes	
Region fixed effects	Yes	
# observations	285	
Adj. R squared	0.9476	
Root MSE	0.0287	

Appendix II

Variance inflation factor (VIF) after regression

Variable	VIF	1/VIF
Rate of employees in the tertiary sector	9.69	0.103226
Rate of informality	7.01	0.142598
Urban population rate	6.89	0.145078
Rate of employees in microenterprises	3.80	0.263328
Gross employment rate	1.84	0.544365
Population with tertiary education	1.70	0.589475
GDP per capita	1.11	0.899874

Appendix III

Breusch and Pagan Lagrangian multiplier test

$EQI[prov,t] = Xb + u[prov] + e[prov,t]$

	Var	Sd=sqrt(Var)
EQI	0.0033794	0.0581323
e	0.0003487	0.0186726
u	0.0002179	0.0147626
Test: Var(u)=0		
chibar2(01)	99.70	
Prob > chibar2	0.0000	

Hausman Test

	Coefficients		(b-B) Difference	Standar error
	(b) Fixed	(B) Random		
Rate of informality	-0.1929	-0.2667	0.0738	0.0189
Population with tertiary education	0.0102	0.0800	-0.0698	0.0148
Interaction between informality and tertiary education	-0.0086	-0.1469	0.1383	0.0307
GDP per capita	-0.0004	0.0004	-0.0008	0.0005
Gross employment rate	-0.1106	-0.0080	-0.1027	0.0185
Rate of employees in microenterprises	-0.2366	-0.1100	-0.1266	0.0306
Urban population rate	0.0138	-0.0554	0.0692	0.0191
Test: H0: difference in coefficients not systematic				
chi2(7)	62.86			
Prob>chi2	0.0000			

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 149-173
https://doi.org/10.38191/iirr-jorr.25.001

Articles

Spatial patterns of crime in Ecuador: analyzing the impact of judicial systems and geographic elements

Carolina Guevara Rosero, Miguel Flores**, Michelle Llumiquinga***, Melany Tulcán*****

Received: 14 July 2024
Accepted: 18 November 2024

Abstract:

This study investigates the spatial patterns of crime in Ecuador and their driving factors, paying special attention to the judicial system's impact. Drawing on data from 218 cantons between 2015 and 2021, we applied exploratory spatial data analysis and spatial econometric models for cross section data and panel data. Our analysis revealed discernible clusters of both high and low crime rates, as well as isolated areas of crime (islands of crime) and safety (islands of non-crime). The findings offer a detailed overview of the crime situation in Ecuador and emphasize the significance of geographic elements in formulating effective crime prevention measures. The analysis also identifies the shift in crime dynamics over time, indicating that cantons typically experiencing low crime rates can evolve into higher crime areas, hinting at a contagion effect within spatial clusters. The study further underscores the critical influence of the judicial system on crime prevalence, where systemic inefficiencies such as case backlogs and a high proportion of unsentenced inmates are associated with rising crime. For policymakers, these insights underscore the necessity of tailoring interventions to the specific contexts and dynamics of each region, considering both the local conditions and the broader surrounding crime environment.

Keywords: Crime; spatial; cluster; Latin America; judicial.
JEL Classification: P25; H50; O17.

Patrones espaciales de la criminalidad en Ecuador: analizando el impacto de los sistemas judiciales y elementos geográficos

Resumen:

Este estudio investiga los patrones espaciales del crimen en Ecuador y sus factores asociados, enfatizando el rol del sistema judicial. Utilizando datos de 218 cantones entre 2015 y 2021, se realiza un análisis exploratorio de datos espaciales y consecuentemente se estiman modelos econométricos espaciales con datos de sección transversal y panel de datos. Nuestro análisis reveló clústeres espaciales tanto de altas tasas de criminalidad como de bajas tasas de criminalidad, así como áreas aisladas (islas de crimen) de criminalidad y áreas aisladas de seguridad (islas de no crimen). Los hallazgos ofrecen una visión detallada de la situación del crimen en Ecuador y enfatizan en la importancia de los elementos geográficos en la formulación de medidas efectivas de prevención del crimen. La investigación también identificó cambios en la dinámica del crimen a lo largo del tiempo, e indica que los cantones que experimentan típicamente

* Departamento de Ciencias Administrativas, Escuela Politécnica Nacional, Quito. Ecuador. carolina.guevara@epn.edu.ec
** Departamento de Economía. Tecnológico de Monterrey. México. miguelflores@tec.mx
*** Escuela Politécnica Nacional, Quito. Ecuador. stefany.llumiquinga@epn.edu.ec
**** Escuela Politécnica Nacional, Quito. Ecuador. melany.tulcan@epn.edu.ec
Corresponding author: miguelflores@tec.mx

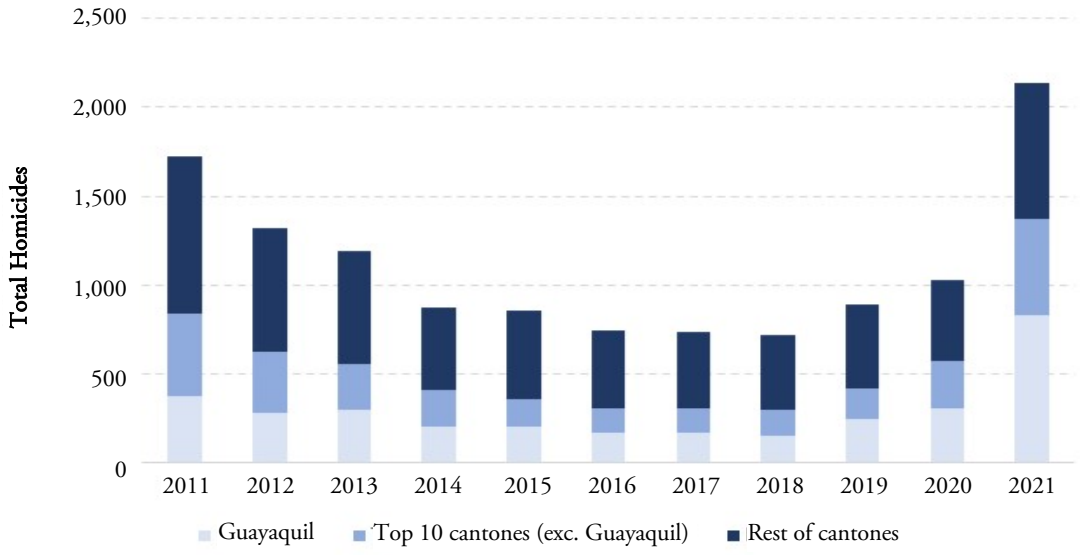
tasas bajas de criminalidad pueden pasar a ser áreas de mayor criminalidad, lo que sugiere un efecto de contagio dentro de los clústeres espaciales. El estudio subraya, además, la influencia del sistema judicial en la prevalencia del crimen, donde las ineficiencias sistémicas como el atraso en los casos y una alta proporción de reclusos sin sentencia se asocian con el aumento del crimen. Para los hacedores de políticas públicas, estos resultados subrayan la necesidad de adaptar las intervenciones a los contextos y dinámicas específicas de cada región, considerando tanto las condiciones locales como el entorno delictivo que lo rodea.

Palabras clave: Crimen; espacial; cluster; América Latina; judicial.
Clasificación JEL: P25; H50; O17.

1. Introduction

According to the United Nations Development Programme-UNDP (2021), the Latin American and the Caribbean region has become the most violent region in the world. Within this context, Ecuador has experienced an increase in violence due to the presence of street gangs, additional to the presence of criminal groups from Colombia, Mexico and Europe for being a strategic point for cocaine trafficking. The increase in homicides in Ecuador has intensified notably since 2018, showing a sustained growth pattern throughout the country. While Guayaquil was initially the epicenter of this violence, the trend has become widespread. As shown in Figure 1, Guayaquil accounted for nearly 39% of all homicides in Ecuador in 2021. The next 10 cantons, excluding Guayaquil, made up 26%, while the remaining cantons across the country accounted for 35%. An increase in violence has been evident in each of these groups of cantons since 2018.

Figure 1.
Homicides Trend in Ecuador by Group of Cantones 2011-2021



By 2021, Ecuador's national homicide rate reached 12.05 per 100,000 inhabitants, making it one of the most dangerous countries in Latin America. This violence is not evenly distributed. Some cantons, primarily along the coast in provinces like Los Ríos, Guayas, Cañar, Esmeraldas, El Oro, and Zamora, recorded over 30 homicides per 100,000 inhabitants. This concentration of violence in specific areas indicates a clear spatial pattern of crime. This may be due to the struggle for territories by important criminal gangs (OECO, 2023). In this context, understanding and addressing crime requires a spatial approach, since crime is not randomly distributed across space (Aguayo & Medellín, 2014; Carrión, 2007; Glaeser et al. 1996) and it diffuses over space (Anselin et al. 2000). Using spatial econometric models,

some studies, considering different types of crime, show that crime is a spatially correlated phenomena, in Colombia (Álvarez and González, 2012; Garza, Nieto and Gutiérrez, 2009), in Mexico (Flores & Villareal, 2015; Nedvedovich, Cervera and Botello, 2017; Flores and Gasca, 2016), in Ecuador (Mena, 2023), in Brazil (Puech, 2004), in El Salvador (Alvarado, 2011), among other countries. Most of them identified clusters of crime but less attention is paid to clusters of non-crime or other types of regions considering their crime dynamics. To address crime, apart from identifying clusters of crime, it is also important to identify regions where crime is not high but are surrounded by regions with high crime levels, and vice versa. In addition, crime has been analyzed considering the intricate interplay of factors like socioeconomic conditions, urban planning, and law enforcement strategies, contributing to specific crime patterns (Graif et al 2014). In developing countries countries, apart from aforementioned factors, the low quality of institutions, encompassing law enforcement, the judiciary and penal system, can make that crime deterrence policies be not effective (Paternoster, 2010).

The goal of this study is to assess the spatial dependence of crime in Ecuador, specifically focusing on violent homicides, and to explore the factors associated with it, with a particular emphasis on the role of the judicial system. The paper conducts a detailed spatial analysis of homicide rates at the canton level, identifying various crime clusters and tracking their evolution over time. This analysis enhances the understanding of violence patterns and helps inform the development of mitigation strategies. Additionally, the study evaluates the impact of the judiciary on violence at the subnational level, providing valuable insights for policymakers. To carry out the study, different data sources are used: the Ministry of the Interior, the Ministry of Education, the National Employment and Unemployment Survey (ENEMDU), the State Attorney General's Office, the Ombudsman's Office, the National Service for Attention to Adults Deprived of Liberty and Adolescent Offenders (SNAI) and the Ecuadorian Observatory of Organized Crime (OECE). Our database accounts for information of 218 cantons from 2015 to 2021. An Exploratory Spatial Data Analysis and spatial econometric models are used.

The rest of the paper is structured as follows: Section 2 reviews existing research on crime theory and the judicial system's role in crime control. Section 3 outlines the data and methodology used in the study. Section 4 presents the results of the analysis, and Section 5 offers concluding remarks.

2. Literature review

According to Becker's seminal model (1968), the decision to engage in criminal behavior is primarily guided by assessments of potential benefits and losses. Within this framework, losses are related legal consequences including imprisonment. The likelihood of facing legal consequences depends on two main factors: police action and the legal capacity of a nation. These factors are complementary. Police actions without the effectiveness of the judicial system will not lead to less crime.

Improvements in legal capacity can act as a deterrent against violent conduct, primarily by altering the perceived risk of imprisonment, heightening the likelihood of punishment or intensifying penalties. The literature emphasizes the vital role of strong judicial systems, judicial independence, reforms, and interventions in reducing crime and creating safer societies (Paternoster 2010). The criminal justice system is crucial for maintaining law and order in society, with the judiciary being a key component responsible for fair dispute resolution, justice administration, and deterring criminal activity (UNODC, 2006). Paternoster (2010) argues that the efficacy of deterrence policies hinges on the quality of institutions, encompassing law enforcement, the judiciary, and the penal system. Therefore, deterrence policies are more effective in nations with superior institutions. Addressing criminality and delinquency requires an effective and efficient justice administration system that enjoys the population's trust in terms of its legitimacy to resolve both coexistence and violence-related issues (Benavides et al. 2016). In a review on the link between judicial corruption and crime rates, Powell et al (2010) indicated that judicial corruption is positively associated with an increase in crime rates, emphasizing the corrosive impact of corruption on the effectiveness of the judicial system.

Kuckertz (2022) takes a unique approach by employing a difference-in-differences analysis to estimate the effects of judicial reform on crime rates in Mexico. The study reveals that judicial reform is associated with a reduction in crime rates, shedding light on the potential of legal reforms in curbing

criminal activity. Apart from judicial capacity and reforms, some studies have explored the impact of various criminal justice interventions on crime rates. Tonry (2010) conducts a comprehensive review of empirical evidence on the effect of pretrial detention on crime. His analysis suggests that pretrial detention may paradoxically lead to an increase in crime rates, highlighting the need for careful consideration of detention policies. Petersilia (2003) reviews empirical evidence on the impact of sentencing reform on crime. Her findings reveal that sentencing reform can have a substantial impact on crime rates, emphasizing the potential of policy changes in shaping criminal behavior. MacKenzie (2006) explores the impact of rehabilitation programs on crime, demonstrating that such programs can effectively reduce recidivism and prevent future crimes. Similarly, Umbreit et al. (2007) finds that restorative justice approaches can be effective in reducing recidivism and fostering community healing.

If the judicial system works, the probability of incarceration increases influencing on criminal behavior. There is substantial evidence indicating that detention serves to reduce criminal offenses, primarily due to an incapacitation effect (Rose and Shem-Tov, 2021). Additionally, Bhuller et al. (2020) demonstrated that incarceration might also contribute to crime reduction after release, as it provides access to social and rehabilitation programs for those who have been convicted.

Regarding the effect of police actions, many studies focus on the impact of policing strategies, such as increasing the number of police officers or implementing hot spot policing interventions. They consistently find that these strategies reduce crime, often without notable spatial displacement effects (Di Tella and Schargrodsky, 2004; Draca et al., 2015; Cheng and Long, 2018; Mello, 2019; Braga et al., 2019).

However, there are exceptions to this trend. For instance, Blattman et al. (2021) discovered that doubling police patrols and enhancing municipal services in Bogotá had a limited impact on crime reduction in targeted areas, while property crimes increased in surrounding neighborhoods. Notably, they found no evidence of violent crimes being displaced. Strategies aimed at reducing police response times can enhance the likelihood of solving crimes and apprehending offenders, which, in turn, contributes to lower crime rates (Blanes, Vidal and Kirchmaier, 2018; Weisburd, 2021). Doleac (2017) examined the expansion of DNA databases of criminal offenders in U.S. states and demonstrated their effectiveness in deterring crime. The use of information technology by police to predict crimes has also been shown to improve crime clearance rates (Mastrobuoni, 2020).

Nevertheless, it's important to recognize that efforts to combat and penalize crime effectively can have unintended consequences contingent on the specific context. Dell (2015) discussed how crackdowns initiated by mayors in Mexico led to a significant increase in homicides as they weakened existing drug traffickers and fueled conflicts between rival traffickers. In alignment with this finding, Acemoglu et al. (2020) revealed that the introduction of high-powered incentives for the Colombian military resulted in a higher number of civilian casualties, particularly in areas with weaker judicial institutions.

In the context of crime, it is important to consider a theoretical aspect called ecological fallacy, which occurs when researchers make incorrect inferences about individual behavior of offenders or victims based on data collected at a group level (Sampson, 2013; Lawton, 2018) such as geographic units like neighborhoods, cities, states, or countries. In the realm of crime studies, the ecological fallacy has been linked to the social disorganization theory, which are complementary to get a better understanding of the relationship between neighborhood characteristics and criminal behavior (Bellair et al. 2019). Social disorganization theory asserts that high crime rates are a consequence of the breakdown of social cohesion within neighborhoods and the inability to regulate behavior effectively. It attributes the disparity across neighborhoods in crime to specific community characteristics or ecological factors such as poverty, unemployment, frequent residential turnover, and racial composition, which can lead to social breakdown, termed as social disorganization. This weakens social order and control mechanisms, contributing to elevated crime levels (Sampson and Groves 1989; Kubrin et al. 2015). Thus, social disorganization theory underscores the significance of neighborhood-level factors in shaping crime patterns. The ecological fallacy serves as a reminder of the complexities in extrapolating individual actions from neighborhood-level data, reinforcing the need for nuanced analyses when applying social disorganization theory to understand and address crime within communities.

Crime is influenced by various economic and sociocultural factors, including population density, unemployment, income, poverty, and education (Mokline, 2018). Numerous studies have shown that

unemployment and poverty have a notable and positive impact on crime rates (Imran et al. 2018; Buonanno et al., 2015; Fadaei-Tehrani & Green, 2002). Based on Becker's economic theory of crime, individuals living in poverty are more likely to engage in property crimes, such as burglary, theft, motor vehicle theft, and arson, because they perceive a greater potential gain from each unlawful act compared to the general population. In poverty conditions, property crimes become a more attractive option for some individuals seeking financial gain (Becker 1968). According to Jawadi et al. (2021), maintaining a stable economy is vital for controlling non-violent crime, as economic downturns lead to increased criminal activity, while economic prosperity reduces it. In addition, crime is a spatial phenomenon with geographical patterns. Glaser et al. (2022) proposed spatial panel models for counts to predict crime counts over space.

Education plays a significant role in reducing crime rates (Huang, Maassen van den Brink, & Groot, 2009), particularly in the case of violent and property crimes (Bradley & Green. 2020). The impact of education on crime can be explained through three main channels in socio-economic literature: income effects, time availability, and differences in patience or risk aversion (Lochner and Moretti 2004). Education tends to decrease criminal participation by increasing the rewards of legitimate work, raising the opportunity costs of illegal activities, and limiting the available time for criminal behavior among teenagers. Additionally, education can influence crime rates by affecting individuals' patience and risk aversion, as those with more patience tend to place higher value on future earnings (Groot et al. 2010; Machin et al. 2011). However, it is important to note that education alone does not guarantee lawful behavior.

3. Data and methodology

3.1. Data

To analyze the factors influencing crime rates in Ecuador, our research utilizes information from the Ministry of the Interior, the Ministry of Education, the National Survey of Employment and Unemployment (ENEMDU), the State Attorney General's Office, the Ombudsman's Office, the National Service for Comprehensive Attention to Adult Persons Deprived of Liberty and Adolescent Offenders (SNAI, acronym in Spanish), and the Ecuadorian Organized Crime Observatory (OECO, acronym in Spanish). Our compiled database encompasses information about 218 cantons from the period of 2011 to 2021. Cantons (3) from the Galápagos province are not considered due to the unavailability of data for them.

The dependent variable in our study is the murder rate, which is quantified as the number of murders per 100,000 inhabitants. Murder is defined as an intentional crime against human life under specific conditions such as premeditation, cruelty, financial incentive, or promise of reward. It is classified as intentional homicide and is subject to penalties as outlined in Article 140 of the Comprehensive Organic Penal Code of Ecuador (Ministry of Justice, 2017). Murder accounts for homicides but is distinguished by its greater criminal intensity. Consequently, for the purposes of this study, the murder rate includes also the crime of contract killing. This indicator serves as a reasonable proxy variable for measuring levels of violence, as endorsed by the United Nations Office on Drugs and Crime (UNODC, 2019). Their accurate recording and comparability across contexts make them valuable for tracking violence trends (Pinker, 2011). Studies demonstrate a strong correlation between homicide rates and other violent crimes, supporting their use as a proxy for overall violence (Eisner, 2003; Zimring & Hawkins, 1997). Furthermore, homicide rates reflect underlying social and economic factors influencing violence, such as inequality and institutional weakness (Fajnzylber et al., 2002; Lederman et al., 2002). Analyzing murder trends provides critical insights for understanding and addressing the complex dynamics of violence. The following Table 1 presents the description of variables and their information sources, while Table A1 in the Appendix shows the descriptive statistics for these variables.

TABLE 1.
Variable description and information sources

Variable name	Description	Information source	Expected result according to the literature
Congestion rate	The congestion rate (CR) is calculated using the formula: CR=(Cn+Cn−1CRn*CR =(Cn+Cn−1CRn*) where: Cn = Cases filed in the current period n in each province Cn-1= Cases filed in the prior period in each province CRn = Cases resolved in the current period in each province	Judiciary Council	(+) Kuckertz (2022); (Paternoster 2010)
Population Density	It is the number of inhabitants per squared kilometer in each canton	National Institute of Statistic and Census (INEC)	(+) (Mokline, 2018)
Employment rate	It is the percentage of employed people over the economically active population in each province.	National Employment, Unemployment and Underemployment Survey – INEC.	(-) Jawadi et al. (2021)
Rate of young people who do not study	Number of young people between 15 and 19 years old who do not study (not enrolled in educational institutions) divided by the total number of young people between 15 and 19 years old	Ministry of Education of Ecuador (MINEDUC)	(+)(Huang, Maassen van den Brink, & Groot, 2009); (Mokline, 2018)
Income Poverty	It is the ratio between the number of people with monthly income lower than 88.72 over the total population in each province.	National Survey of Employment, Unemployment and Underemployment – INEC.	(+) (Imran et al. 2018; Buonanno et al., 2015; Fadaei-Tehrani & Green, 2002)
Unsentenced prison population index	It is the ratio between the number of people deprived of liberty who have not been sentenced (processed) and the total number of people deprived of liberty.	The National Service for Comprehensive Attention to Adult Persons Deprived of Liberty and Adolescent Offenders (SNAI)	(+) Tonry (2010)

TABLE 1. CONT.
Variable description and information sources

Variable name	Description	Information source	Expected result according to the literature
Seized Weapons per inhabitant	It is the number of weapons seized by the National Police per 100 thousand inhabitants in each province.	Ecuadorian Observatory of Organized Crime (OECO)	(-)
Serious Complaints	It is the ratio of serious complaints—such as illegal money collection, smuggling, organized crime, extortion, money laundering, illegal production of controlled substances, kidnappings for ransom, planting or cultivation of substances, illegal possession and carrying of weapons, illicit trafficking of firearms and substances, and human trafficking—relative to the population of each canton.	Ecuadorian Observatory of Organized Crime (OECO)	(-)
Drug seizure	It is the quantity of illegal controlled substances (in kilograms) seized by the National Police intended for domestic consumption per 1000 inhabitants in each province.	Ecuadorian Observatory of Organized Crime (OECO)	(+)

3.2. METHOD

For the analysis of crime in Ecuador, two models were estimated. The first model (Model A) is estimated using a spatial panel data model, incorporating the following independent variables: congestion rate, rate of young people who do not study, employment rate, poverty rate, index of unsentenced prison population and population density. The time frame covered for this model is from 2015 to 2021. The second model (Model B) is estimated using a cross-sectional spatial model, with the independent variables from the first model, plus additional provincial-level variables: seized weapons reported, reports of serious crime complaints, and drug seizure. Population density is excluded in the cross-section model due to multicollinearity. This model specifically analyzes the year 2021 due to data availability.

a) Spatial panel model

Spatial panel data consists of observations that are included in two dimensions: cross-sectional and temporal. This model allows addressing the unobserved heterogeneity. For this study, the panel data used

for Model A consists of *i* spatial observations (218 cantons) and time period *t* (6 years), resulting in a balanced panel. Furthermore, in a spatial panel framework, either a fixed-effects model or a random-effects model can be applied. To determine which model to use, the Hausman test is conducted. Among the diversity of spatial models, the spatial lag model (SLM) is used in the present study. This is a model where the spatial parameter is determined using the lagged spatial dependent variable, Wy, as an additional explanatory variable, and is specified as follows:

$$y_{it} = \rho \sum_{j=1}^{N} W_{ij} y_{jt} + x_{it}\beta + \mu_i + \varepsilon_{it} \quad (1)$$

where i is an index for the cross-sectional dimension (cantons), with =1;...;218; t is an index for the temporal dimension (years), with t=1;...;6; $\boldsymbol{y_{it}}$ represents the murder rate for each canton i in year t; $\boldsymbol{W_{ij}}$ is an element of the spatial weights matrix of dimension N x N; $\boldsymbol{x_{it}}$ is the matrix of dimension NT x k of observations of the independent variables, where k is the number of independent variables (see Table 1); ρ is the spatial parameter associated with the dependent variable; β is the vector of unknown parameters corresponding to the independent variables. The term $\boldsymbol{\mu_i}$ is the fixed spatial effect that captures the unobserved heterogeneity produced by variables that vary across cantons but remain constant over time; $\boldsymbol{\varepsilon_{it}}$ is a vector of independently and identically distributed error terms, capturing the unobserved heterogeneity produced by variables that change both over time and across cantons. For the model estimations, two weight matrices are used: i. the inverse distance weight matrix and ii. the 4 nearest neighbor matrix[1]. The latter was chosen based on the R squared of the estimations between the dependent variable and the spatial lag dependent variable, using different number of neighbors. In panel data analysis, the Hausman test allows for choosing between a fixed-effects model or a random-effects model. According to the results of the Hausman test (Chi-squared=35.773, p-value=0.000008001), the preferred model is the fixed-effects model.

b) Cross-sectional spatial model

Spatial regression models in a cross-section setting allow for examining spatial dependence (spatial correlation) and spatial heterogeneity among observations.

$$y = \rho W y + X\beta + \mu \quad (2)$$

Where y represents the murder rate for each canton i in the year 2021; W is the spatial weight matrix; ρ is the spatial autoregressive parameter; X is an $n \times K$ matrix of observations of the independent variables (see Table 1); β is the vector of unknown parameters corresponding to the observations of the independent variables.

To select between the spatial lag model and the spatial error model, two metrics are used: the Lagrange multiplier test for spatial dependence in linear model test (Anselin, Bera, & Florax, 1996; Borrego Sánchez, 2018) and the Akaike's Information Criterion and the Bayesian Information Criterion (Agiakloglou, C., & Tsimpanos, A. 2023). Table 2 shows that according to the LM test, the spatial lag model is the preferred model in comparison to the spatial error model. Table 3 indicates that the preferred models are the spatial lag models since the AIC and BIC criteria are lower for these models than for the spatial error models.

[1] The rationale behind using these two matrices was to ensure that our analysis captured the spatial relationships between cantons in different ways, reflecting both proximity-based and neighbor-based spatial dependencies. Despite the differences in how these matrices define spatial relationships, the results from both approaches were remarkably consistent, which strengthens the robustness of our findings. By obtaining similar results across both matrices, we can confidently assert that the observed spatial patterns of crime and its determinants are not dependent on the specific choice of the spatial weight matrix. This robustness supports the validity of our spatial econometric models and ensures that the spatial correlations and the effects of the judicial system, socioeconomic variables, and other factors on crime rates are reliable and not an artifact of any particular spatial structure.

TABLE 2.
LM test for spatial models

Test	LM for the panel data model (inverse distance)	LM- for the panel data model (Kn4)	LM- for the cross section model (inverse distance)	LM for the cross section model (Kn4)
LM_{LAG} (p-valor)	9.1419 (0.002498)	19.069 (0.00001261)	1.215 (0.2703)	3.0592 (0.08028)
LM_{ERROR} (p-valor)	5.8185 (0.01586)	17.392 (0.00003042)	0.12802 (0.7205)	1.814 (0.178)

TABLE 3.
AIC and BIC statistics for spatial models

Test	AIC for the panel data model (inverse distance)	BIC- for the panel data model (inverse distance)	AIC for the panel data model (Kn4)	BIC- for the panel data model (Kn4)	AIC for the cross section model (inverse distance)	BIC for the cross section model (inverse distance)	AIC for the cross section model (Kn4)	BIC for the cross section model (Kn4)
$Model_{LAG}$	1572.985	1606.83	1571.365	1605.21	1575.345	1615.958	1573.747	1614.361
$Model_{ERROR}$	1573.636	1607.481	1572.241	1606.086	1576.42	1617.034	1574.786	1615.4

4. RESULTS

Our results stem from the spatial analysis of crime, as showed in Figure 1 and Table 4 and from the econometric models showed in Table 5. For the spatial analysis of crime, the LISA indicator of local spatial correlation is applied using two weight matrices: i. the inverse distance weight matrix (Figure 1a) and ii. the 4 nearest neighbor-based matrix (Figure 1b). The results indicate that there exist clusters of crime and non-crime, islands of crime and islands of non-crime. Overall, they are not constant over time, which reflects that the phenomenon of crime is dynamic, changing according to the time and geographical context. Table 4 details the number and names of cantons in each cluster, based on the two weight matrices, in three specific years: 2011, 2016 and 2021.

a) Clusters of crime

In 2011, Cluster 1, as depicted in Figures 1a and 1b, consisted of 9 cantons using the inverse distance matrix and 7 cantons using the 4kn weight matrix. Over time, this cluster has shown a tendency to shrink. By 2016 and 2021, Cluster 1 included fewer cantons. An interesting observation is that Atacames and Río Verde remained part of this cluster throughout the years, suggesting their significance as potential hubs for organized crime.

In 2021, Cluster 2 comprised 19 cantons when using the inverse distance matrix and 7 cantons when using the 4kn weight matrix. This cluster significantly reduced in size by 2016. Notably, 8 out of the 24 cantons in the 2021 cluster (using the inverse distance matrix) were consistent with the 2011 cluster. These cantons are El Empalme, Balzar, Palestina, Daule, Samborondón, Mocache, Pueblo Viejo, and Babahoyo, indicating the presence of well-established criminal networks. In 2021, Cluster 2 reappeared in a more southern location, consisting of 8 cantons using the 4kn weight matrix.

A third crime cluster was identified in 2011. Using the inverse distance matrix, this cluster was composed of Cuyabeno (Sucumbíos), and using the 4kn weight matrix, it was composed of Putumayo (Sucumbíos). Although different cantons were identified, they are neighboring regions near the Colombian border. In 2016, the size of this cluster grew, but it completely disappeared by 2021 according to both matrices.

b) Non-crime clusters

In 2011, Cluster 4 of non-crime, using the inverse distance matrix, included 73 cantons primarily from the Azuay and Loja provinces. Analyzing with the 4kn weight matrix, this cluster consisted of 22 cantons, mainly from Azuay. By 2016, the cluster's size, using the inverse distance matrix, reduced to 20 cantons, retaining 18 of the original ones. However, with the 4kn weight matrix, the cluster expanded to 36 cantons in 2016, keeping 7 original cantons.

For 2021, the size of the non-crime cluster increased to 52 cantons using the inverse distance matrix, incorporating 18 new cantons. Conversely, with the 4kn weight matrix, the cluster size decreased to 32 cantons in 2021, also including 18 new cantons. Notably, 19 cantons remained consistent within this non-crime cluster, regardless of the matrix used. It is worth mentioning that Quito was part of this non-crime cluster in both 2011 and 2021.

a) Islands of crime

These spatial clusters refer to cantons with high rates of violent deaths per inhabitant surrounded by cantons with low rates. In 2011, using both weight matrices, Celica (Loja) consistently stands out as an island of crime. Other islands of crime are identified depending on the matrix used. By 2016, using both matrices, Pablo VI (Morona Santiago), Paquisha, Palanda (Zamora Chinchipe), Macará, and Paltas (Loja) emerged as consistent islands of crime. In 2021, Nangaritza (Zamora Chinchipe) consistently stands out as an island of crime using both matrices. These islands require attention because interventions to reduce crime in these areas might be effective due to the apparent lack of a strong territorial presence of criminal bands around them.

b) Islands of non-crime

These clusters refer to cantons with low rates of violent deaths per inhabitant surrounded by cantons with high rates. In 2011, using the inverse distance matrix, there were 12 islands of no crime. When using the 4kn weight matrix, this cluster consisted of 2 cantons. The cantons that appeared in both analyses were Puerto Quito and Pedro Vicente Maldonado (Pichincha). By 2016, using the inverse distance matrix, there was 1 island of no crime, while the 4kn weight matrix showed 4 cantons in this cluster. The canton repeated in both analyses was Muisne (Esmeraldas). In 2021, the inverse distance matrix identified 16 islands of no crime, whereas the 4kn weight matrix showed a cluster of 8 cantons. Special attention should be given to these islands of non-crime to prevent the diffusion of crime from neighboring cantons.

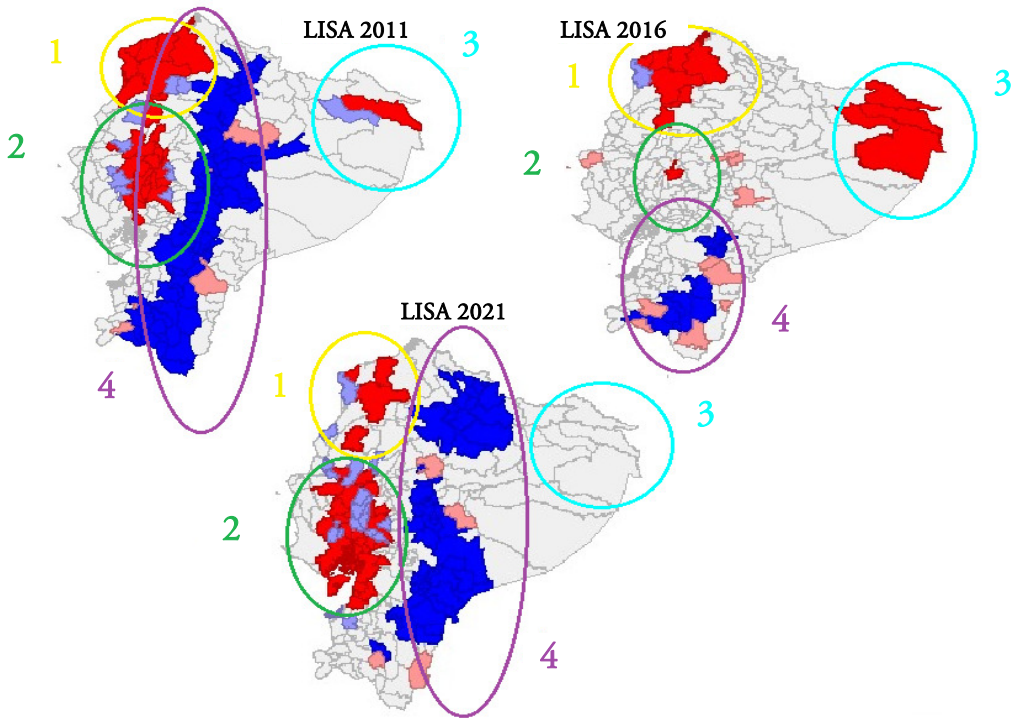
c) Changing cantons

According to the changing status of cantons, four types of cantons can be identified: i. Black sheep becoming white sheep among black sheep: These are cantons with high values of violent deaths that are surrounded by other cantons with high values, which eventually change to low values of violent deaths. ii. Black sheep becoming white sheep among white sheep: These are cantons with high values of violent deaths that are surrounded by cantons with low values, which eventually change to low values of violent deaths. iii. White sheep becoming black sheep among white sheep: These are cantons with low values of violent deaths that are surrounded by cantons with low values, which eventually change to high values of violent deaths. iv. White sheep becoming black sheep among black sheep: These are cantons with low values of violent deaths that are surrounded by cantons with high values, which eventually change to high values of violent deaths.

A detailed description of each type of canton is provided in Appendix C.

FIGURE 2.
Clusters of crime according to the LISA indicator

a) LISA indicator using inverse distance weight matrix



b) LISA indicator using the 4 nearest neighbor weight matrix

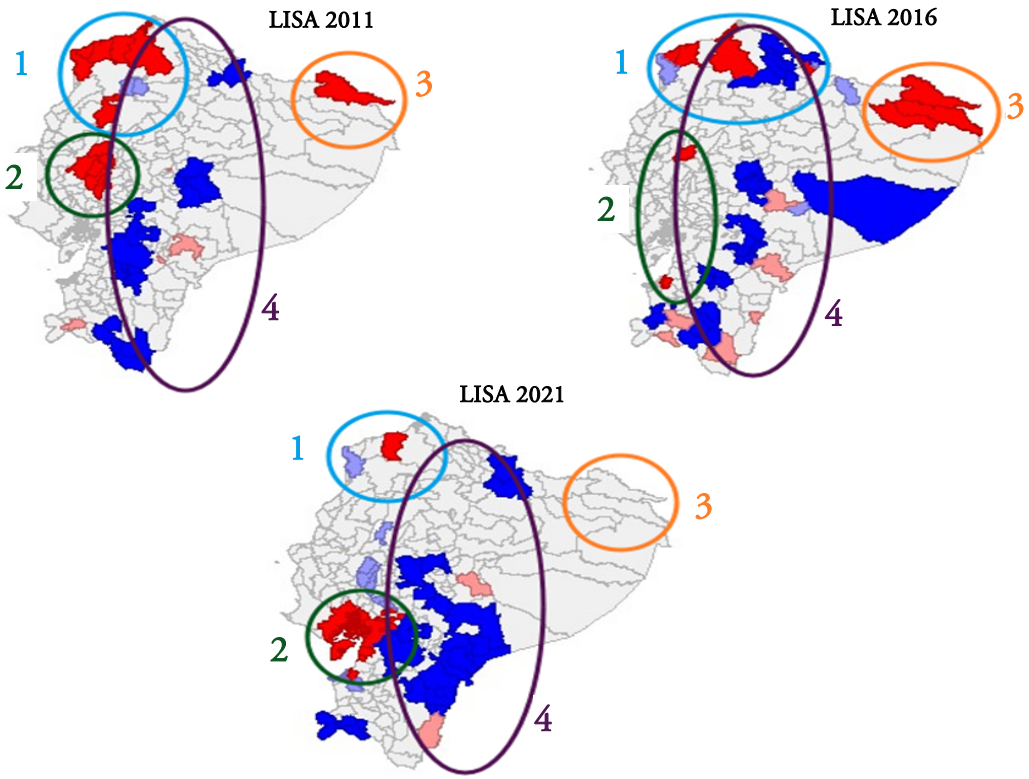


Table 4.
Composition of clusters of crime, non-crime, islands of crime and non-crime

Cluster type	Year 2011		Year 2016		Year 2021	
	4kn	Inverse distance	4kn	Inverse distance	4kn	Inverse distance
High-High 1	Muisne, Atacames, Esmeraldas, Río Verde, Eloy Alfaro (Esmeraldas), La Concordia (Santo Domingo de los Tsáchilas), El Carmen (Manabí).		Eloy Alfaro, Esmeraldas, Atacames, (Esmeraldas)		Río Verde (Esmeraldas)	
	7 cantons:	9 cantons:	4 cantons:	8 cantons:	1 canton:	4 cantons:
		Quinindé (Esmeraldas), Pedernales (Manabí).	Montúfar (Carchi)	Río Verde, Quinindé (Esmeraldas), La Concordia (Santo Domingo de los Tsáchilas), Puerto Quito (Pichincha), El Carmen (Manabí).		Quinindé, Atacames (Esmeraldas), Flavio Alfaro (Manabí).
High-High 2	El Empalme, Balzar (Guayas), Mocache, Palenque, Vinces, Baba, Pueblo Viejo (Los Ríos)				Guayaquil, Naranjito, Naranjal, Balao (Guayas), La Troncal (Cañar)	
	7 cantons:	19 cantons:	2 cantons:	1 canton:	8 cantons:	24 cantons:
		Daule, Durán, Palestina, Salitre, Samborondón (Guayas), Pichincha (Manabí), San Jacinto de Buena Fe, Quevedo, Quinsaloma, Ventanas, Babahoyo, Urdaneta (Los Ríos).	El Empalme (Guayas), Machala (El Oro)	Ventanas (Los Ríos)	Durán, El Triunfo (Guayas), Machala (El Oro)	Santa Ana, Bolívar (Manabí), Colimes, Balzar, El Empalme, Palestina, Santa Lucía, Pedro Carbo, Daule, Samborondón, Yaguachi, Milagro, Lomas de Sargentillo (Guayas), Mocache, Pueblo Viejo, Urdaneta, Babahoyo (Los Ríos), El Guabo (El Oro), Camilo Ponce Enríquez (Azuay).

TABLE 4. CONT.
Composition of clusters of crime, non-crime, islands of crime and non-crime

Cluster type	Year 2011		Year 2016		Year 2021	
	4kn	Inverse distance	4kn	Inverse distance	4kn	Inverse distance
High-High 3			Putumayo, Shushufindi, Cuyabeno (Sucumbíos)			
	1 canton:	1 canton:	3 cantons:	4 cantons:		
	Putumayo (Sucumbíos).	Cuyabeno (Sucumbíos).		Aguarico (Orellana).		
Low-Low 4	Pimampiro (Imbabura), Baños, Tisaleo (Tungurahua), Penipe, Pallatanga (Chimborazo), Palora, Pablo VI (Morona Santiago), Cañar, Suscal (Cañar), Cuenca, San Fernando, Girón, Nabón, Oña (Azuay), Calvas, Espíndola, Gonzanamá (Loja)		Azogues, Biblián (Cañar), Gualaceo, El Pan, Guachapala, Paute, Sevilla de Oro (Azuay), Calvas, Quilanga, Gonzanamá, Loja, Catamayo (Loja)		Sucumbíos, Gonzalo Pizarro (Sucumbíos), San Pedro de Huaca (Carchi), Ambato, Pelileo (Tungurahua), Colta, Guamote (Chimborazo), Morona, Logroño, Tiwintza, Santiago de Méndez, Limón Indanza, San Juan Bosco, Gualaquiza (Morona Santiago), Chordeleg, Oña (Azuay), El Pangui, Yacuambi, Yantzaza (Zamora Chichipe)	
	22 cantons:	73 cantons:	36 cantons:	20 cantons:	32 cantons:	52 cantons:
	Sucumbíos (Sucumbíos), Chillanes (Bolívar), Cumandá (Chimborazo), Santa Isabel (Azuay), Palanda, Chinchipe (Zamora Chinchipe)	Montúfar, Bolívar (Carchi), Otavalo, Ibarra, Antonio Ante (Imbabura), Quito, Cayambe, Mejía, Pedro Moncayo, Rumiñahui (Pichincha), Latacunga, Salcedo, Saquisilí (Cotopaxi), Píllaro, Ambato, Mocha, Quero, Pelileo, Patate (Tungurahua), Mera, Santa Clara, (Pastaza), Carlos Julio Arosemena, Tena	Tulcán, Mira (Carchi), Cotacachi, Ibarra, San Miguel de Urcuquí (Imbabura), Pastaza (Pastaza), Ambato, Mocha, Tisaleo, Cevallos, Quero, Pelileo (Tungurahua), Guano, Penipe, Riobamba, Alausí (Chimborazo),	Déleg (Cañar), Chordeleg, Oña (Azuay), Celica, Sozoranga, Olmedo (Loja), Yacuambi, Zamora (Zamora Chinchipe).	Montúfar (Carchi), Guaranda, Chimbo (Bolívar), Cuenca, Santa Isabel (Azuay), Centinela del Cóndor, Paquisha (Zamora Chinchipe), Zapotillo, Celica, Macará, Sozoranga (Loja), Cañar, Suscal (Cañar)	El Ángel, Bolívar, San Gabriel (Carchi), Ibarra, Antonio Ante, Otavalo, Pimampiro (Imbabura), El Chaco, Quijos (Napo), Cayambe, Quito, Pedro Moncayo (Pichincha) Saquisilí, San Miguel de Salcedo (Cotopaxi), Tisaleo, Cevallos, Mocha, Quero (Tungurahua), Riobamba, Penipe, Guano, Chambo (Chimborazo), Pablo Sexto, Sucúa (Morona

TABLE 4. CONT.
Composition of clusters of crime, non-crime, islands of crime and non-crime

Cluster type	Year 2011		Year 2016		Year 2021	
	4kn	Inverse distance	4kn	Inverse distance	4kn	Inverse distance
		(Napo), Huamboya (Morona Santiago), Guano, Riobamba, Chambo, Colta, Alausí, Chunchi (Chimborazo), Biblián, El Tambo, Azogues, Déleg (Cañar), Paute, Sevilla de Oro, Guachapala, El Pan, Gualaceo, Sigsig (Azuay), Saraguro, Loja, Chaguarpamba, Catamayo, Paltas, Sozoranga, Macará, Olmedo, Quilanga (Loja), Yacuambi, Zamora, Palanda, Chinchipe (Zamora Chinchipe), Portovelo, Piñas (El Oro).				Santiago), Azogues (Cañar), Nabón, Sevilla de Oro, Gualaceo, El Pan, Paute, Sigsig, Guachapala (Azuay), Catamayo (Loja).
Low-High Islands of non-crime	Puerto Quito, Pedro Vicente Maldonado (Pichincha)		Muisne (Esmeraldas)		Muisne (Esmeraldas), Buena Fe, Palenque, Vinces, Baba (Los Ríos), Alfredo Baquerizo Moreno, Simón Bolívar (Guayas), Santa Rosa (El Oro)	
	2 cantons:	12 cantons:	4 cantons:	1 canton:	8 cantons:	16 cantons:
		Flavio Alfaro, Bolívar, Olmedo (Manabí), Shushufindi (Sucumbíos), Colimes,	San Pedro de Huaca (Carchi), Cascales (Sucumbíos),			Jama, Pichincha, Tosagua, Junín (Manabí), Salitre, Isidro Ayora, Nobol, General Antonio Elizalde (Guayas)

TABLE 4. CONT.
Composition of clusters of crime, non-crime, islands of crime and non-crime

Cluster type	Year 2011		Year 2016		Year 2021	
	4kn	Inverse distance	4kn	Inverse distance	4kn	Inverse distance
		Santa Lucía, Lomas de Sargentillo (Guayas), Las Naves, Echeandía (Bolívar), Montalvo (Los Ríos).	Huamboya (Morona Santiago)			
High-Low Islands of crime	Celica (Loja)		Pablo VI (Morona Santiago), Paquisha, Palanda (Zamora Chinchipe), Macará, Paltas (Loja)		Nangaritza (Zamora Chinchipe)	
	4 cantons:	4 cantons:	6 cantons:	10 cantons:	2 cantons:	4 cantons:
	Cevallos (Tungurahua), Santiago de Méndez (Morona Santiago), Chordeleg (Azuay)	Balsas (El Oro), Gualaquiza (Morona Santiago), Cevallos (Tungurahua).	Limón Indanza (Morona Santiago)	Montecristi (Manabí), Gualaquiza (Morona Santiago), Sigsig (Azuay), Píllaro (Tungurahua), San Miguel de Salcedo (Cotopaxi).	Palora (Morona Santiago)	Latacunga (Cotopaxi), Gonzanamá (Loja), Pastaza (Pastaza).

Table 5 presents the results of our models: the panel data spatial model in column (1) considering data for cantons from 2011 to 2021, and the cross-section spatial model in column (2) considering data for cantons for 2021. The spatial panel model accounts for the unobservable heterogeneity of cantons. However, due to limited availability of temporal information for some variables, this model includes fewer variables than the cross-section spatial model. Additional variables such as seized weapons per inhabitant, serious complaints, and drug seizures were included in the cross-section spatial model, thereby reducing the omitted variable bias.

Overall, the parameter of spatial correlation is highly significant, indicating that an increase in the murder rate in neighboring cantons leads to an increase in the murder rate in a given canton. The dimension of the spatial correlation is higher when using the inverse distance matrix compared to the 4kn weight matrix. This demonstrates that crime is a spatial phenomenon, concentrating in specific regions and then diffusing to neighboring areas. The analysis of local spatial correlation indicated that there are clusters of crime in the northern and southern parts of the coastal region and along the border with Colombia.

Our variables of interest related to the judicial system include the cause congestion rate, the unsentenced prison population, and the efficiency of the judicial system. The first two variables are significant and positive, indicating a positive association with the level of crime. According to the social disorganization theory (Bellair, 2017) and considerations of ecological fallacy, these factors can be seen as contextual-level factors that shape the spatial crime pattern. A 1% increase in the cause congestion rate leads to a 2.84% increase in the murder rate. This indicates that when the judicial system does not resolve cases timely, the level of crime increases. The inefficiency of the judicial system can be perceived by criminals and associated with a lower probability of being punished, thereby inducing more crime. According to a jurisprudence study by Enríquez Burbano (2017), Ecuador's traditional judicial system does not allow for the use of alternative means for justice, which augments the cause congestion rate. Furthermore, the system has insufficient economic resources to resolve all judicial cases.

Similarly, an increase in the non-sentenced prison population correlates with higher crime levels. A 1% increase in this population leads to an approximate 0.08% increase in the murder rate. This suggests that despite police actions to capture criminals, the collapse of the judicial system leads to more crime for two reasons: first, not all arrested criminals are sentenced, allowing complaints to lapse and criminals to be released to continue their activities; second, a high rate of unsentenced arrested criminals signals judicial inefficiency, further augmenting crime rates. Contrary to findings by Rose and Shem-Tov (2021) for North Carolina, detention itself does not reduce criminal offenses in Ecuador, despite the incapacitation effect. In Ecuador, criminal gang operations primarily occur in prisons, allowing gang members to maintain contact with their leaders and reinforce criminal activities. This undermines the expected inverse relationship between imprisonment and crime as noted by Withers (1984). While Bhuller et al. (2020) argued that incarceration might reduce crime post-release due to social and rehabilitation programs, this is not the case in Ecuador where prison conditions are inadequate for social rehabilitation. The efficiency of the judicial system, measured as the ratio between the prison population and the number of complaints, has no effect on the murder rate, possibly due to the presence of other variables that better proxy judicial system efficiency.

In the judicial system context, the number of serious complaints (per 100,000 inhabitants) related to illegal activities such as smuggling, organized crime, extortion, and money laundering is associated with a reduction in the violent murder rate. This result indicates that reporting crimes can reduce crime levels, as authorities can intervene and legal processes can begin to determine punishments or sanctions. Londoño & Guerrero (1999) also found that the propensity to report reduces the violence rate. However, this propensity could decrease due to judicial system inefficiency (Carrión, 2022). There is a positive association between drug seizures and the murder rate. A 1% increase in seized drugs per inhabitant is associated with 0.88 more violent murders per inhabitant. This can be explained by the fact that police drug seizure operations occur in significant distribution and operational centers for drug dealers (Bulla et al., 2016) where violent acts are more common. Additionally, drug seizure operations often lead to confrontations between the police and cartel members, causing deaths.

TABLE 5.
Results of the model estimations of violent deaths rate

Variables	Panel data model (inverse distance matrix)	Cross-section model (inverse distance matrix)	Panel data model (4kn matrix)	Cross-section mode (4kn matrix)
Congestion rate	3.823*** (0.794)	7.134* (2.321)	3.788*** (0.795)	6.957** (2.326)
Rate of young people who do not study	0.080* (0.039)	0.008 (0.087)	0.078* (0.039)	0.008 (0.086)
Employment rate	-0.339** (0.118)	-0.264 (0.409)	-0.322** (0.118)	-0.234 (0.407)
Income poverty	0.267*** (0.077)	0.027 (0.296)	0.262*** (0.076)	0.027 (0.295)
Income poverty 2	-0.003*** (0.001)	-0.001 (0.003)	-0.003** (0.001)	-0.001 (0.003)
Unsentenced prison population index	0.049*** (0.012)	0.139** (0.047)	0.049*** (0.012)	0.135** (0.047)
Population Density	0.0006 (0.0005)		0.0007 (0.0005)	
Seized Weapons per inhabitant		0.069 (0.068)		0.070 (0.068)
Serious Complaints		-0.236** (0.113)		-0.236* (0.112)
Drug seizure		1.062* (0.582)		1.054* (0.579)
ρ (Rho)	0.504*** (0.119)	0.459 (0.304)	0.160*** (0.038)	0.152 (0.089)

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, standard errors in parentheses.

Regarding the socioeconomic situation of regions, poverty, employment, and the rate of young people not in education are important factors influencing crime levels. In Ecuador, the non-monotonic relationship between poverty and violent deaths holds true, as Sánchez and Nuñez (2001) found for Colombia. Crime increases with poverty up to a certain point, where it then starts to decrease. In Ecuador, the average violent death rate per 100,000 inhabitants is 4.98 for the poorest cantons (first quartile), 9.11 for the second quartile, 8.24 for the third quartile, and 5.45 for the wealthiest cantons (fourth quartile). Crime increases up to 36% poverty, beyond which it starts to decrease. In the poorest cantons, crime levels are lower than in moderately poor cantons. This non-linear relationship might be explained by the theory of relative poverty (Braithwaite, 1979). In highly impoverished areas, inequality is evident, potentially leading poor individuals to commit crimes against wealthier individuals to compensate for perceived disparities (Fajnzylber, Lederman, & Loayza, 2002). In the most impoverished areas, where everyone is equally poor, this inequality—and consequently crime—decreases. People in poverty can be more vulnerable, influenceable, and prone to entering the crime environment. According to Becker's (1968) theory, individuals turn to crime when potential benefits outweigh the costs, making those living in poverty more likely to engage in criminal activities. Additionally, in the context of poverty, social bonds may be

weak, impeding informal social controls and lowering the likelihood of reporting general deviance to authorities (Sampson, 2013).

Regarding the employment rate, a 1% increase is associated with a 0.52% decrease in violent deaths per 100,000 inhabitants. When a region offers job opportunities, people have a stable economy, which discourages participation in criminal activities. Therefore, as noted by Sani et al. (2018), Jawadi et al. (2021), and Fadaei-Tehrani & Green (2002), a stable economy with job opportunities and economic prosperity is vital for reducing crime and violence. This is exemplified by Esmeraldas, which registers a low level of employment (89.98%) with respect to the national average (95.88%).

Lack of education, measured by the rate of young people who are not studying, is positively associated with the crime rate. It has been observed that criminal gangs involved in drug trafficking engage young people in their activities. These young individuals are more likely not to be in school, giving them free time and a low opportunity cost for engaging in crime. The school dropout rate was also tested in the models, but it was not significant in any of them. This result indicates that what drives crime is the high number of young people who do not engage in education at all, rather than those who initially participate in the education system and later drop out. Contrary to expectations, population density and judicial system efficiency were fund not statistical significant factors in explaining crime rates in the present analysis. This could be related to the presence of criminal gangs in smaller cities in Ecuador. Not all cities with high murder rates have high population densities. For instance, Pueblo Viejo, Naranjal, and Nagaritza have population densities of 134.96, 56.23, and 4.11, respectively, which are lower than the national average of 137.61 inhabitants per square kilometer.

5. Conclusions

This study examined recent crime trends in Ecuador from 2015 to 2021, focusing on the spatial dependence of violent murders at the cantonal level. Utilizing exploratory spatial data analysis and spatial econometric models, the research identified clusters of high and low crime rates as well as isolated clusters of crime and non-crime. These findings offer a comprehensive view of the crime landscape in Ecuador, emphasizing the importance of geographic factors in developing more effective crime prevention policies and strategies.

Additionally, the study explored the evolution of these spatial clusters over time, revealing that the crime status of cantons—whether low-crime "white sheep" or high-crime "black sheep"—can change, indicating a contagion effect. This is particularly notable in cases where a low-crime canton surrounded by high-crime areas shifts towards higher crime rates. For policymakers, these insights underscore the necessity of tailoring interventions to the specific contexts and dynamics of each region, considering both local conditions and the broader surrounding crime environment. This targeted approach could be crucial in effectively addressing and mitigating crime in Ecuador.

Our findings further establish the crucial role of the judicial system in influencing crime rates in Ecuador. Inefficiencies within this system, indicated by high cause congestion rates and a substantial unsentenced prison population, correlate with an increase in crime. The slow processing of cases might signal to potential offenders that the likelihood of punishment is low, thereby reducing the deterrent effect of the law. Additionally, the high unsentenced prison population highlights a disconnection between the police and the judicial system, where the courts fail to promptly sentence individuals apprehended by law enforcement. Moreover, with gang leaders operating within prisons, incarceration may inadvertently serve as a reward rather than a deterrent, providing criminals with opportunities to network and strengthen gang ties rather than being rehabilitated.

Contrary to expectations, our analysis shows that population density does not significantly impact crime rates across Ecuadorian regions. This phenomenon may be attributed to the presence of gangs in smaller, economically disadvantaged areas characterized by low employment and moderately high poverty rates. Given these insights, several policy implications emerge. First, it is vital to address the geographic patterns of crime and consider the specific background and surrounding environment of each region. The evident judicial congestion necessitates innovative reforms to enhance efficiency in the judicial sector.

However, merely improving judicial processes is insufficient. A comprehensive reform must also include restructuring and eliminating corruption from prisons to transform them into facilities that truly aim for social rehabilitation.

Furthermore, eradicating poverty is imperative. Investing in education and economic opportunities can decrease the allure of criminal activities, particularly among the youth. These multi-faceted approaches are essential for developing effective strategies to reduce crime in Ecuador, highlighting the need for targeted interventions that address both the symptoms and root causes of criminal behavior.

Regarding limitations, our study faced challenges due to the limited availability of temporal data for all variables. To address this, we estimated two models: a cross-sectional spatial econometric model and a panel spatial model. While our analysis primarily related crime to the murder rate, focusing solely on intentional homicide, it is important to acknowledge that crime, particularly organized crime, encompasses a broader spectrum of criminal activities. These include property robberies, terrorist acts, and extortion targeting residents and businesses. This complexity suggests that our findings, while insightful, might not fully capture the multifaceted nature of crime.

References

Acemoglu, D., Fergusson, L., Robinson, J., Romero, D., & Vargas, J. F. (2020). The perils of high-powered incentives: evidence from Colombia's false positives. *American Economic Journal: Economic Policy*, *12*(3), 1-43.

Agiakloglou, C., & Tsimpanos, A. (2023). Evaluating the performance of AIC and BIC for selecting spatial econometric models. *Journal of Spatial Econometrics*, *4*(1), 2.

Aguayo Téllez, E., & Medellín Mendoza, S. E. (2014). Dependencia espacial de la delincuencia en Monterrey, México. *Ecos de Economía*, *18*(38), 63-92. Obtenido de http://www.scielo.org.co/pdf/ecos/v18n38/v18n38a3.pdf

Alvarado, W. R. (2011). Modelo Geoestatico espacio-temporal del crimen en el Salvador: análisis Estructural y predictivo. *Revista de Matematica: Teoria y Aplicaciones*, *2*(18), 325-342. Obtenido de https://www.scielo.sa.cr/pdf/rmta/v18n2/a07v18n2

Anselin, L., Cohen, J., Cook, D., Gorr, W. & Tita, G. (2000). Spatial Analyses of Crime. In D. Duffee (ed.), *Criminal Justice 2000: Volume 4. Measurement and Analysis of Crime and Justice.* National Institute of Justice.

Anselin, L., Bera, A. K., & Florax, R. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77-104. https://doi.org/10.1016/0166-0462(95)02111-6

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of political economy*, *76*(2), 169-217.

Bellair, P. (2017). Social disorganization theory. In *Oxford Research Encyclopedia of Criminology and Criminal Justice*.

Bellair, P. E., McNulty, T. L., & Piquero, A. R. (2019). Persistent material hardship and childhood physical aggression. *Aggression and violent behavior*, *49*, 101309.

Benavides, F., Binder, A., Villadiego, C., & Niño, C. (2016). La reforma a la justicia en América Latina: las lecciones aprendidas. *Bogotá, Colombia: Friedrich-Ebert-Stiftung en Colombia.*

Blanes i Vidal, J., & Kirchmaier, T. (2018). The effect of police response time on crime clearance rates. *The Review of Economic Studies*, *85*(2), 855-891.

Blattman, C., Duncan, G., Lessing, B., & Tobón, S. (2021). *Gang rule: Understanding and countering criminal governance* (No. w28458). National Bureau of Economic Research.

Bradley, S., & Green, C. (Eds.). (2020). *The economics of education: A comprehensive overview.* Elsevire, Academic Press.

Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: An update of an ongoing systematic review and meta-analysis. *Journal of experimental criminology*, *15*, 289-311.

Braithwaite, J. (1979). *Inequality, crime and public policy.* Routledge and Kegan Paul.

Bhuller, M., Dahl, G. B., Løken, K. V., & Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, *128*(4), 1269-1324.

Bulla, P., Ramírez, B., García, J. F., Lovera, M. P., & Wiesner, D. (2016). *Homicidios y venta de drogas: una peligrosa dupla en Bogotá.* Fundación ideas para la Paz.

Buonanno, P., et al. (2015). Poor Institutions, Rich Mines: Resource Curse in the Origins of the Sicilian Mafia. *Economic Journal, 125*(586), F175-F202.

Carrión, F. (2007). La geografía del delito. *Ciudad Segura: Programa estudios de la ciudad* 1(17). https://repositorio.flacsoandes.edu.ec/bitstream/10469/2642/1/BFLACSO-CS17-01-Carri%c3%b3n.pdf

Carrión, F. (2022). La violencia en el Ecuador, una tendencia previsible. *Ecuador Debate* (117), 15-40. Obtenido de http://hdl.handle.net/10469/19199

Cheng, C., & Long, W. (2018). *The Effect of Highly Publicized Police-Related Deaths on Policing and Crime: Evidence from Large US Cities*. Working Paper.

Di Tella, R., & Schargrodsky, E. (2004). Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, *94*(1), 115-133.

Dell, M. (2015). Trafficking networks and the Mexican drug war. *American Economic Review*, *105*(6), 1738-1779.

Doleac, J. L. (2017). The effects of DNA databases on crime. *American Economic Journal: Applied Economics*, *9*(1), 165-201.

Draca, M., & Machin, S. (2015). Crime and economic incentives. *Economics*, *7*(1), 389-408.

Eisner, M. (2003). Long-term historical trends in violent crime. *Crime and Justice*, *30*(1), 83-142.

Enríquez Burbano, G. (2017). El procedimiento abreviado como una forma de descongestión del sistema judicial penal. *Revista de la Facultad de Jurisprudencia*, (2), 1-37. https://doi.org/10.26807/rfj.v1i2.24

Fadaei-Tehrani, R. & Green, T. M. (2002). Crime and society. *International Journal of Social Economics*, *29*(10), 781-795.

Fajnzylber, O., Ledereman, D., & Loayza, N. (2002). Inequality and violent crime. *The Journal of Law and Economic, 45*(1), 1-40

Flores, M., & Gasca, F. (2016). Patrones de distribución espaciotemporal de los homicidios de mujeres en México en 1990, 2000 y 2010. *Sociedad y economía,* (32), 15-40.

Flores, M., & Villarreal, A. (2015). Exploring the Spatial Diffusion of Homicides in Mexican Municipalities Through Exploratory Spatial Data Analysis. *A Journal of Policy Development and Research*, *17*(1), 35-50. Obtenido de http://www.jstor.org/stable/26326920

Garza, N., Nieto, L., & Gutiérrez, M. (2009). El homicidio en Barranquilla: una lectura espacial. *Revista de Economía del Caribe*, (3), 176-229.

Glaeser, E., Sacerdote, B. and Scheinkman, J. (1996). Crime and Social Interactions. *Quarterly Journal of Economics*, *111*(2), 507-548.

Glaser, S., Jung, R. C., & Schweikert, K. (2022). Spatial panel count data: modeling and forecasting of urban crimes. *Journal of Spatial Econometrics*, *3*(1), 2.

Graif, C., Gladfelter, A. S., & Matthews, S. A. (2014). Urban poverty and neighborhood effects on crime: Incorporating spatial and network perspectives. *Sociology compass*, *8*(9), 1140-1155.

Groot, W., & van den Brink, H. M. (2010). The effects of education on crime. *Applied economics*, *42*(3), 279-289.

Huang, J., van den Brink, H. M., & Groot, W. (2009). A Meta-Analysis of the Effect of Education on Social Capital. *Economics of Education Review, 28*(4), 454-464.

Imran, M., Hosen, M., & Chowdhury, M. A. F. (2018). Does Poverty Lead to Crime? Evidence from the United States of America. *International Journal of Social Economics, 45*(10), 1424-1438.

Jawadi, F., Mallick, S. K., Cheffou, A. I., & Augustine, A. (2021). Does higher unemployment lead to greater criminality? Revisiting the debate over the business cycle. *Journal of Economic Behavior & Organization*, *182*, 448-471.

Kubrin, C. E., & Wo, J. C. (2015). Social disorganization theory's greatest challenge: linking structural characteristics to crime in socially disorganized communities. *The handbook of criminological theory*, 121-136.

Kuckertz, R. E. (2021). Torture, Mistreatment, and Forced Confessions in Mexico's Accusatorial Criminal Justice System. *Mexican law review*, *14*(1), 3-51.

Lawton, B. (2018). Deciding on the "Appropriate" Unit of Analysis. In *The Oxford Handbook of Environmental Criminology*. Oxford University Press.

Lederman, D., Loayza, N. V., & Menendez, A. M. (2002). Violent crime: Does social capital matter? *Economic Development and Cultural Change*, *50*(3), 509-539.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, *94*(1), 155-189.

Londoño, J., & Guerrero, R. (1999). Violencia en América Latina Epidemiología y Costos. Banco Interamericano de Desarrollo. Red de Centros de Investigación de la Oficina del Economista Jefe. https://www2.utp.edu.co/~porlapaz/docs/violencia/pxp2.pdf

Machin, S., Marie, O., & Vujić, S. (2011). The crime reducing effect of education. *The Economic Journal*, *121*(552), 463-484.

MacKenzie, D. L. (2006). *What works in corrections: reducing the criminal activities of offenders and deliquents*. Cambridge University Press.

Mastrobuoni, G. (2020). Crime is terribly revealing: Information technology and police productivity. *The Review of Economic Studies*, *87*(6), 2727-2753.

Mello, S. (2019). More COPS, less crime. *Journal of public economics*, *172*, 174-200.

Mastrobuoni, G. (2020). Crime is terribly revealing: Information technology and police productivity. *The Review of Economic Studies*, *87*(6), 2727-2753.

Ministry of Justice. (2017). Código Orgánico Integral Penal. https://www.oas.org/juridico/PDFs/mesicic5_ecu_ane_con_judi_c%C3%B3d_org_int_pen.pdf

Mokline, A. (2018). Criminality, Human Capital and Economic Performance. *Business and Economic Research*, *8*(4), 130-142.

Mena Quinteros, D. A. (2023). *Factores determinantes de las actividades delictivas en el Ecuador: Un análisis econométrico espacial.* [undergraduate thesis.] Universidad Tecnica de Ambato.

Nedvedovich, S. S., Cervera Gómez, L. E., & Botello Mares, A. (2017). México: territorialización de los homicidios. Las razones de la violencia en el norte del país. Realidad, datos y espacio. *Revista internacional de estadística y geografía, 8*(2). https://rde.inegi.org.mx/rde_21/RDE21.pdf

OECO. (2023b). *Evaluación situacional del entorno estratégico del narcotráfico en Ecuador 2019-2022*. Observatorio Ecuatoriano del Crimen Organizado. https://oeco.padf.org/wp-

content/uploads/2023/07/EVALUACION-SITUACIONAL-NARCOTRAFICO-ECU-2019-2022-.pdf

Paternoster, R. (2010). How much do we really know about criminal deterrence. *J. crim. l. & criminology*, *100*, 765.

Petersilia, J. (2003). *When prisoners come home: Parole and prisoner reentry*. Oxford University Press.

Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. Penguin.

Powell, B., Manish, G. P., & Nair, M. (2010). 13 Corruption, crime and economic growth. *Handbook on the Economics of Crime*, 328.

Puech, F. (2004, December). How do criminals locate? Crime and spatial dependence in Minas Gerais. In *Documento presentado en la conferencia International Economic Policies in the New Millenium, Coimbra, Portugal*. Disponible en: http://www4.fe.uc.pt/30years/papers/68.pdf

Rose, E., & Shem-Tov, Y. (2021). How does incarceration affect crime? Estimating the dose-response function. *Journal of Political Economy*.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American journal of sociology*, *94*(4), 774-802.

Sampson, R. J., & Wilson, W. J. (2013). Toward a theory of race, crime, and urban inequality. In *Race, crime, and justice* (pp. 177-189). Routledge.

Sánchez Torres, F. J., & Núñez Méndez, J. A. (2001). Determinantes del crimen violento en un país altamente violento: el caso de Colombia. *Economía, crimen y conflicto*, 285.

Umbreit, M. S., Coates, R. B., & Vos, B. (2007). Restorative justice dialogue: A multi-dimensional, evidence-based practice theory. *Contemporary Justice Review*, *10*(1), 23-41.

UNODC. (2006). *The Integrity and Accountability of the Police: Criminal Justice Assessment Toolkit*.

UNODC. (2019). *Estudio Mundial sobre el homicidio. Oficina de las Naciones Unidas contra la Droga y el Delito*. https://www.unodc.org/documents/ropan/2021/HOMICIOS_EN_ESPANOL.pdf

UNDP, United Nations Development Program. (2021). *Informe regional de desarrollo humano 2021, Atrapados: alta desigualdad y bajo crecimiento en América Latina y el Caribe.* Centro Regional de Servicios para América Latina y el Caribe, Dirección Regional para América Latina y el Caribe, Programa de las Naciones Unidas para el Desarrollo (PNUD). Obtenido de https://www.undp.org/es/el-salvador/publications/informe-regional-sobre-desarrollo-humano-2021

Tonry, M. (2010). Public criminology and evidence-based policy. *Criminology & Pub. Pol'y*, *9*, 783.

Weisburd, S. (2021). Police presence, rapid response rates, and crime prevention. *Review of Economics and Statistics*, *103*(2), 280-293.

Withers, G. (1984). Crime, Punishment and Deterrence in Australia: An Empirical Investigation. *Economic Record, 60*(2), 176-185. https://doi.org/10.1111/j.1475-4932.1984.tb00851.x

Wortley, R., Sidebottom, A., Tilley, N., & Laycock, G. (Eds.). (2018). *Routledge handbook of crime science.* Routledge.

Zimring, F. E., & Hawkins, G. (1997). *Crime is not the problem: Lethal violence in America*. Oxford University Press.

ORCID

Carolina Guevara Rosero https://orcid.org/0000-0001-7605-1443

Miguel Flores https://orcid.org/0000-0001-8660-219X

Michelle Llumiquinga https://orcid.org/0009-0005-6114-5881

Melany Tulcán https://orcid.org/0009-0002-7768-989X

Appendixes

Appendix A

Table A1.
Descriptive statistics of the independent variables of year 2021

Variable	Mean	Standard deviation	Min	Max
Congestion rate	1.806	0.3470599	1.402	2.524
Rate of young people who do not study	47.843	8.14523864	16.45	69.06
Employment rate	95.88	2.276573	89.23	98.93
Income poverty	34.013	13.6021633	17.616	70.508
Unsentenced prison population index	29.166	16.760	0	65.854
Population Density	137.606	371.656	0.311	4786.789
Serious Complaints	12.776	15.5151	0.442	73.208
Weapons seized/provincial population per 1000 inhabitants	10.3902	10.3845	0.333	36.673
Drug seizures (kg)/provincial population per 1000 inhabitants	1.2466	2.7491	0.047	16.34

Appendix A2

From the temporal analysis, 4 kinds of cantons could be identified: i. black sheep becoming white sheep among black sheep; that is cantons with high values of violent deaths surrounded by other cantons with high values that change to low values of violent deaths; ii. black sheep becoming white sheep among white sheep, that is cantons with high values of violent deaths surrounded by cantons with low values that change to low values of violent deaths; iii. white sheep becoming black sheep among white sheep, that is cantons with low values of violent deaths surrounded by cantons with low values that change to high values of violent deaths; and iv. white sheep becoming black sheep among black sheep, that is cantons with low values of violent deaths surrounded by cantons with high values that change to high values of violent deaths.

Using both matrices, Muisne canton is identified as black sheep in 2011 (recorded 16.80 violent deaths per 100000 inhab.) becoming white sheep among black sheep in 2016 (recorded 3.26 violent deaths per 100000 inhab.) and 2021 (recorded 6.42 violent deaths per 100000 inhab.).

Regarding the second kind of changing cantons, using both matrices, Célica (Loja) stands as black sheep in 2011 (recorded 13.14 violent deaths per 100000 inhab.), becoming a white sheep in 2016 among white sheep (recorded 0 violent deaths per 100000 inhab.). Using the inverse distance weight matrix, Gualaquiza (Morona Santiago) was a black sheep in 2011 (22.08 violent deaths per 100000 inhab.) and

2016 (5.25 violent deaths per 100000 inhab.) and became a white sheep in 2021 (0 violent deaths per 100000 inhab.), influenced by an environment of white sheep. Canton Cevallos (Tungurahua) recorded a high level of crime in 2011 (11.62 violent deaths per 100000 inhab.), became a White sheep in 2021 (0 violent deaths per 100000 inhab.).

Using the 4kn weight matrix, Chordeleg (Azuay) was a black sheep in 2011 (15.05 violent deaths per 100000 inhab.) and became a white sheep in 2021 ((0 violent deaths per 100000 inhab.), influenced by a white sheep environment. Pillaro (Tungurahua) was also a black sheep in 2016 (4.76 violent deaths per 100000 inhab.) and became a white sheep in 2021 (2.29 violent deaths per 100000 inhab.), influenced by a white sheep environment.

As for the third kind of changing cantons, using the inverse distance matrix, Píllaro (Tungurahua), Salcedo (Cotopaxi), Palanda (Zamora Chinchipe), Sigsig (Azuay), Macará, Paltas (Loja), Gonzanamá (Loja) and Pablo Sexto (Morona Santiago) registered low levels of crime in 2011 and in 2016 and became black sheep in 2021 despite of being surrounded by an environment of white sheep. Using 4kn weight matrix, Nangaritza (Zamora Chinchipe) was a white sheep in 2016 and became a black sheep in 2021. Sucumbíos was a white sheep in 2011 and became black sheep in 2016, surrounded by a white sheep environment.

Using the inverse distance weight matrix, group 4 of changing cantons include Puerto Quito (Pichincha) and Shushufinfi (Sucumbíos), which registered low levels of crime in 2011 (0 and 8.56 violent deaths per 100000 inhab., respectively) and became black sheep in 2016 (12.79 and 11.33 violent deaths per 100000 inhab.), influenced by an environment of black sheep. Using the 4kn weight matrix, Milagro (Guayas) was a white sheep in 2011 and 2016 (6.84 and 3.17 violent deaths per 100000 inhab., respectively) and became a black sheep in 2021 (23.22 violent deaths per 100000 inhab.), influenced by an environment of black sheep.

TABLE A3.

Variables	Panel data model (inverse distance matrix)	Cross-section model (inverse distance matrix)	Panel data model (4kn matrix)	Cross-section mode (4kn matrix)
Congestion rate	3.833*** (0.795)	7.317 (2.200)	3.798*** (0.796)	7.123 (2.212)
Rate of young people who do not study	0.081* (0.039)	0.025 (0.077)	0.079* (0.039)	0.023 (0.076)
Employment rate	-0.337** (0.118)	-0.245 (0.403)	-0.320** (0.118)	-0.214 (0.401)
Income poverty	0.266*** (0.077)	-0.020 (0.284)	0.263*** (0.076)	-0.021 (0.282)
Income poverty 2	-0.003** (0.001)	-0.001 (0.003)	-0.003** (0.001)	-0.001 (0.003)
Unsentenced prison population index	0.050*** (0.012)	0.139 (0.045)	0.049*** (0.012)	0.135 (0.045)
Population Density	0.001 (0.001)		0.001 (0.001)	
Seized Weapons per inhabitant		0.075 (0.070)		0.076 (0.070)
Serious Complaints		-0.248 (0.108)		-0.249 (0.108)

TABLE A3. CONT.

Variables	Panel data model (inverse distance matrix)	Cross-section model (inverse distance matrix)	Panel data model (4kn matrix)	Cross-section mode (4kn matrix)
Drug seizure		1.115 (0.525)		1.110 (0.523)
Capital	0.173 (0.642)	4.220 (1.863)	0.194 (0.640)	4.317 (1.853)
ρ (Rho)	0.508*** (0.119)	0.493 (0.291)	0.161*** (0.038)	0.164 (0.088)
AIC	1571.435	1572.28	1569.595	1570.403
BIC	1608.665	1616.279	1606.825	1614.402

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, standard errors in parentheses.

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 175-195
https://doi.org/10.38191/iirr-jorr.25.002

Articles

An empirical examination of the dynamics of tourism, growth, and economic development at a regional level. The colombian departments

Juan Gabriel Brida, Pablo Juan Cárdenas-García**, María Leivy Mejía-Alzate***, Verónica Segarra*****

Received: 09 May 2024
Accepted: 04 November 2024

Abstract:

Since the normalization of tourism in Colombia resulting from improved security and the pacification of various regions of the country, increased attention has been paid to the key role of tourism in regional economic development. The objective of this paper is to analyze the relationship between the tourism sector and economic development in Colombia. To examine the dynamic relationship existing between international tourism and economic development, the 33 departments of Colombia are analyzed between 2012 and 2021. The empirical analysis is conducted using non-parametric tools, derived from symbolic time series analysis, based on the notion of economic regime and clustering techniques. This offers a contribution in two aspects: on the one hand, a novel methodology is applied, and, on the other hand, regional-level analysis is performed. A limited mobility of the departments between the different regimes has been observed during this period, indicating the stability of the variables analyzed. The results of the study suggest that the relationship between tourism and economic development differs in the different groups of departments identified, offering evidence of differences within the country. It should be noted that the group of departments with the highest level of tourism displays a high level of economic development and growth.

Keywords: Tourism; economic growth; economic development; regional analysis; dynamic regime.
JEL Classification: R11; O10.

Un examen empírico de la dinámica entre turismo, crecimiento y desarrollo económico a nivel regional en los departamentos colombianos

Resumen:

Desde la normalización del turismo en Colombia, resultado de la mejora de la seguridad y la pacificación de varias regiones del país, se ha prestado mayor atención al papel clave del turismo en el desarrollo económico regional. El objetivo de este trabajo es analizar la relación entre el sector turístico y el desarrollo económico en Colombia. Para examinar la relación dinámica existente entre el turismo internacional y el desarrollo económico, se analizan los 33 departamentos de Colombia entre 2012 y 2021. El análisis empírico se realiza utilizando herramientas no paramétricas, derivadas del análisis simbólico de series temporales, basadas en la noción de régimen económico y técnicas de clustering.

* GIDE - Facultad de Ciencias Económicas y de Administración, Universidad de La República. Uruguay. gabriel.brida@fcea.edu.uy
** Departamento de Economía. Universidad de Jaén. España. pcgarcia@ujaen.es
*** Facultad de Administración. Institución Universitaria Colegio Mayor de Antioquia. Colombia. maria.mejia@colmayor.edu.co
**** GIDE - Facultad de Ciencias Económicas y de Administración. Universidad de La República. Uruguay. veronica.segarra@fcea.edu.uy
Corresponding author: pcgarcia@ujaen.es

Esto supone una contribución en dos aspectos: por un lado, se aplica una metodología novedosa y, por otro, se realiza un análisis a nivel regional. Se observa una limitada movilidad de los departamentos entre los distintos regímenes durante este periodo, lo que indica la estabilidad de las variables analizadas. Los resultados del estudio sugieren que la relación entre turismo y desarrollo económico difiere en los distintos grupos de departamentos identificados, ofreciendo evidencia de diferencias dentro del país. Cabe destacar que el grupo de departamentos con mayor nivel de turismo presenta un alto nivel de desarrollo y crecimiento económico.

Palabras clave: Turismo; crecimiento económico; desarrollo económico; análisis regional; régimen dinámico.
Clasificación JEL: R11; O10.

1. Introduction

Over recent decades, tourism has become one of the fastest-growing economic sectors worldwide (Meng-Yi Tai & Chao, 2023). In 2022, this economic activity represented 7.6% of the global gross domestic product (7.7 trillion dollars), 9.0% of the global work positions (295 million), and spending by international travelers reached 1.7 trillion dollars, making up 6.8% of the total exportations (World Travel & Tourism Council, 2023).

Tourism has proven to be a major generator of economic growth, creating new jobs, increasing income, stimulating infrastructure development, promoting foreign exchange, and providing tax revenue for the public sector or leveraging other economic areas (Brida et al., 2016; Alcalá-Ordoñez et al., 2023). In addition to tourism's importance as a contributor to economic growth, distinct international organizations (UNWTO, 2018; World Bank, 2019) and scientific studies (Cárdenas-García et al., 2019; Wong et al., 2023), have defended its capacity as a tool to increase the levels of economic development and improve the population's living conditions.

In fact, economic growth and economic development are distinct concepts that are not necessarily related, since greater economic growth does not necessarily imply higher levels of development (Croes et al., 2021). However, economic growth and the income generated may be used to develop policies to improve the population's living conditions (Cárdenas-García et al., 2019). Although initially the increase in Gross Domestic Product (GDP) was considered the traditional objective of public policies, linking macroeconomic objectives to an exclusively monetary measure (Todaro & Smith, 2020), over recent decades, a change in focus has taken place, in which the economic growth perspective is not used exclusively. This new focus includes new objectives that are not necessarily economic in nature, such as education, poverty, health, or inequality. Together with the economic criteria, they may serve as a foundation to measure a country's prosperity in a broader manner (Lee, 2017).

In the case of Colombia, tourism is considered to be a strategic economic activity with the potential to promote economic growth and development. Since 2005, this region has experienced sustained growth (Zapata-Aguirre et al., 2020), with tourism becoming one of the country's main economic activities, after the energy mining sector, and before coffee (Garavito et al., 2019). However, some major differences exist within the country as a whole (Rodríguez-Benavides & Ceballos-Mina, 2022).

Regarding the characterization of tourism, Colombia is a country that is divided politically into 33 departments, each of which is considered to be a region having a unique combination of natural, cultural, and historical resources and attractions that represents a diversification of the tourist offer. The arrival of tourist flows to each of these departments and/or regions, has been an uneven process, given that, except in the case of small island states and microstates, the development of tourism occurs at a regional level (Calero & Turner, 2020). Furthermore, these departments display major differences in terms of their level of progress (Sanabria-Gómez, 2017) or socio-economic development (Galvis-Aponte & Meisel-Roca, 2011).

Although tourism has emerged as an important source of income and improvement of development conditions, little empirical evidence exists on the regional impacts of tourism in Colombia (Sanguinet et al., 2020). In fact, although it has been shown that this sector may contribute approximately one-third of

the growth of the country's overall GDP (Brida, et al., 2017), very little empirical evidence exists on the impact of tourism in the regions of Colombia. The few works that do exist tend to rely on qualitative measures, as is the case with Hernández-García (2013) and Rozo-Bellon & Garavito-González (2014), or they analyze other aspects of tourism activity such as competitiveness (Rodríguez-Victoria et al., 2017).

Therefore, it is necessary to understand and measure the impact of tourism activity in Colombia, taking a regional approach, since tourism has become a strategic sector. Tourism is an activity that impacts destinations. Therefore, it is important to move away from a national to a regional level. If data were available, it would also be ideal to consider the tourist destination level. Indeed, except in the case of small island states or micro-states, the impact takes place with major differences at a regional level (Calero & Turner, 2020). Thus, the contribution of tourism from an economic point of view occurs regionally (Bassil et al., 2023), and this is the most appropriate focus of analysis in an empirical analysis.

This justifies the selection of this country and the approach taken for the case study. No relevant empirical work has analyzed the relationship between tourism, economic growth, and development in Colombia, using a regional approach.

Given the gap in the scientific literature regarding the relationship between tourism, economic growth, and economic development with a regional focus, and the lack of relevant studies considering Colombia as a case study, the objective of this article is to detect homogeneous groups of regions having similar dynamic behavior in terms of tourism behavior, economic growth, and level of economic development. It relies on data from the 33 departments of Colombia during the 2012-2021 period and considers the diversity of regions in terms of tourism specialization, growth dynamics, and economic development.

To characterize tourism activity, information is obtained from the Colombian Ministry of Commerce, Industry and Tourism on the arrival of foreign tourists to each of the 33 departments in absolute terms. Growth dynamics are measured with Gross Domestic Product (GDP) at constant 2015 prices, measured in millions of pesos, relativized in per capita terms. Information was obtained from the Colombian National Administrative Department of Statistics (DANE). Finally, the Human Development Index (HDI) is used to characterize economic development, using data provided by the United Nations Development Programme.

For this dynamic analysis, the notion of economic regime was introduced (Brida et al., 2020), and symbolic time series were used (Risso, 2018). This methodology allows for the grouping of regions that have similar behavior, identifying the characteristics of tourism demand and supply associated with greater economic growth and a higher level of economic development.

The work has been organized as follows: the next section provides a review of the main contributions to the literature on the relationship between tourism, economic growth, and economic development, with a special emphasis on the case of Colombia; the third section justifies the selection of Colombia for the case study and indicates the measurement variables used; the fourth section details the methodology used and presents the results; and finally, the last section offers some final conclusions and recommendations for tourism policy.

2. Tourism, economic growth and economic development

2.1. A general overview

Given the importance of the impacts derived from the expansion of tourism, specifically, the economic impact of this activity, its influence on economic growth began to be analyzed at the beginning of this century. The pioneer study by Balaguer and Cantavella-Jorda (2002) was the starting point of what is today a global research phenomenon, the study of the Tourism-led Growth Hypothesis (TLGH). Following this initial work on the TLGH, numerous authors have conducted empirical analyses on the relationship between tourism and economic growth, determining that, in general, a one-directional relationship exists, from tourism to economic growth (Brida et al., 2016; Alcalá-Ordóñez et al. 2023).

Following this research line, a second current has arisen that analyzes the relationship between tourism and economic development, whereby tourism may be considered an economic activity that permits the improvement of the living conditions of the host population (Cárdenas et al., 2015). As occurred with the study of the relationship between tourism and economic growth, distinct works have begun to appear in the scientific literature, analyzing the relationship between these variables, although not as extensively as in the prior current. The majority of these studies have found a unidirectional causality from tourism to economic development (Alcalá-Ordoñez & Segarra, 2023).

However, in both of these research streams, a major gap exists in the scientific literature, with few works taking a regional level analysis approach, despite the fact that tourism development rarely occurs uniformly across a country.

2.2. Tourism and economic growth in Colombia

Various authors have examined the contribution of tourism to the economic growth of Colombia, generally at a national level, and, to a lesser extent, at a regional or city-wide level.

On the one hand, distinct works have taken a country-level approach; for example, Such-Devesa et al. (2009) quantified the contribution of the tourism sector to the growth of the Colombian economy for the 1990-2006 period. They complemented this analysis with a study of causal relationships between GDP per capita, tourism expenditure in Colombia, and the real exchange rate, drawing conclusions about the importance of the tourism sector on the long-term growth of Colombia's economy. Brida et al. (2009) measured the contribution of tourism to Colombia's economic growth using the Growth Decomposition methodology. They compared their results with those of other surrounding countries such as Argentina, Brazil, Mexico, and Uruguay. Rozo Bellón & Garavito-González (2014) descriptively examined the relationship between tourism and development in Colombia for the 1996-2012 period, concluding that Colombia was experiencing a period in which tourism was beginning to prosper and suggesting that planning would be essential to ensure the success of its tourism policies. Brida et al. (2017) expanded on the growth decomposition methodology using data from the Colombian Tourism Satellite Account to measure the effects of each of the characteristic branches of the tourism sector on Colombian economic growth. Delgado-Munévar (2023) established the relationships existing between economic growth and tourism development in five Latin American countries: Colombia, Ecuador, Guatemala, Mexico, and Peru.

Other works have taken a regional and local approach to the analysis. For example, Brida et al. (2010) measured tourism's contribution to economic growth in the department of Antioquia (Colombia). Brida & Monterubbianesi (2010) analyzed the role of tourism and its effects on economic growth, focusing on four major tourism regions of Colombia. Brida et al. (2020) measured tourism's contribution to economic growth in the city of Medellín for the 2005-2015 period. Sanguinet et al. (2020) used the inter-regional model of input-output to analyze the short-term impact of internal tourism in all of the country's regions, arguing that it may be an efficient mechanism to reduce disparities.

All these studies have revealed the substantial contribution of tourism to the economic growth of Colombia. Indeed, various analyses have shown that tourism plays a significant role at both national and regional levels. Its impact on GDP has been measured, causal relationships with key economic variables have been explored, and its effects have been compared with those of other countries in the region. Furthermore, specific research has been conducted on its influence in distinct regions and cities of Colombia, highlighting its ability to promote local economic growth.

These findings highlight the importance of tourism as an engine of economic growth and suggest its potential as a tool to reduce regional disparities.

2.3. Regional development in Colombia

Various approaches and research areas have been used to study regional economic development in Colombia, although the scientific literature has yet to consider the importance of tourism for development processes.

Some empirical works have revealed the difference in regional development between Bogota and the other departments of the country. For example, Bonet & Meisel (2009) examined regional economic disparities in Colombia by analyzing regional income figures of the country's departments during the 1975-2000 period. They revealed the polarization process arising between Bogota and the other regions since the per capita of this department almost doubles the national average and is more than eight times that of the departments having the lowest income (Choco). Galvis-Aponte et al. (2017) performed an exploratory analysis of the regional figures of economic production in Colombia during the 1975-2015 period, with the results revealing that Bogota has maintained a share of approximately 30% throughout the analyzed years. This figure almost doubles the participation of Antioquia and even exceeds the total of the 23 smallest departments of the country, highlighting the importance and persistence of Bogota as an aggregate value generator at a national level. It also reveals how most of the municipalities having a greater economic richness are situated in the central part of the country.

On the other hand, other empirical work groups have highlighted the existence of factors such as efficiency or geography as being key to the development between regions. For example, Cotte-Poveda (2011) analyzed the economic development and growth in Colombian departments between 1993 and 2007, using the Data Envelopment Analysis (DEA). They used distinct approaches such as poverty, inequality, or security to measure the efficiency of each of the regions, finding that the highest levels of economic activity, quality of life, employment, and security were related to greater efficiency. Peiró-Palomino et al. (2021) offered a composite indicator of well-being for the 33 Colombian departments in 2016, adapting the OECD's Better Life Index at a regional level, including key dimensions such as income, health, education, security, housing, environment, labor market, and civic participation and governance, and applying Data Envelopment Analysis and Multicriteria Decision Making as a methodology. This permitted the comparison of well-being between departments and the creation of rankings, highlighting major disparities in well-being between departments, with population concentration in those with higher levels of well-being. Thus, geography emerged as a relevant factor, revealing a center-periphery duality.

These works highlight the need for policies that address regional disparities and promote more equitable and sustainable development throughout the country. Furthermore, it should also be noted that no research stream addresses the relationship between tourism and regional development in Colombia. This suggests a gap in the research that may be key to an improved understanding of the engines of economic development in different regions of the country.

Therefore, these findings highlight the importance of analyzing the importance of tourism as a tool for economic development.

3. Methodological approach

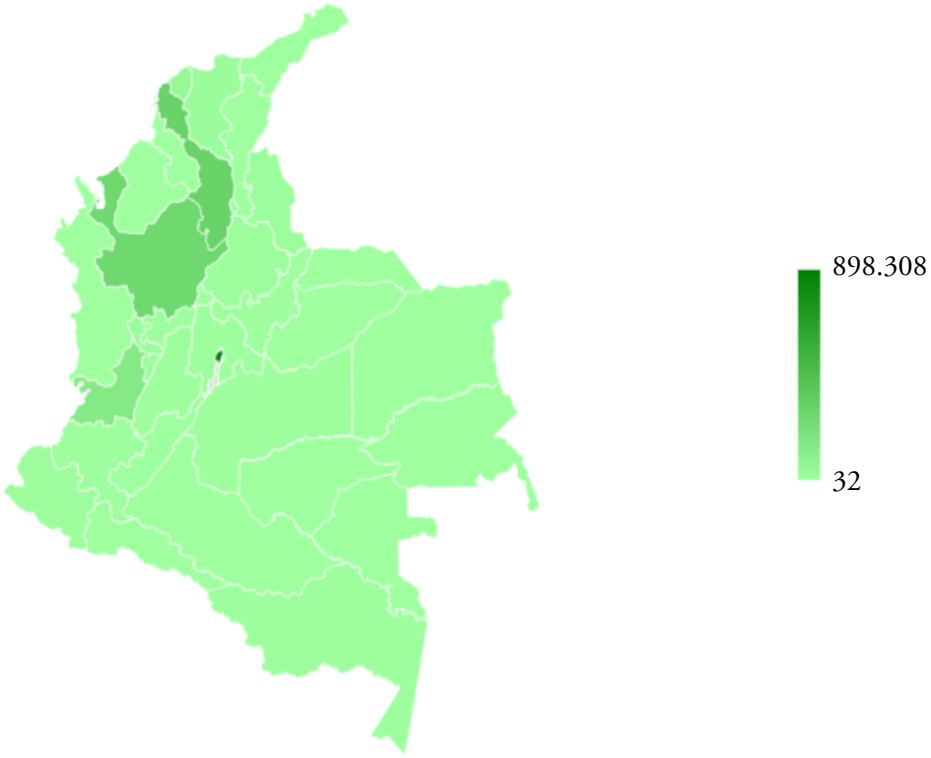
3.1. A regional case study approach

In Colombia, tourism has been established as a strategic economic activity. In 2022, this country received $7.379 billion USD in foreign currency, the highest figure in its history and surpassing pre-pandemic figures (Office of Economic Studies, 2024).

In 2022, tourism's direct contribution to the Colombian GDP was 31,769,400 million COP (2.3% of the GDP), while the total contribution of tourism to the GDP (including the broader effects of investment, supply chain, and induced impacts on income) was 63,357,300 million COP in this same year (4.5% of the GDP) (World Travel & Tourism Council, 2023).

Compared to other economic activities, tourism ranks second after the mining industry, more specifically, the coal and oil sectors (Garavito & Ochoa, 2016). However, despite these data at a national level, there are major differences in tourism's importance in the 33 departments of Colombia, with the following being the most touristic: Bogota, Bolivar, Antioquia, and Valle del Cauca. Here, the average data on international tourists exceeds 100,000 tourists per year. These four departments, on average, received approximately 85% of the country's international tourists during the 2012-2021 period (Figure 1). The following map of Colombia shows its 33 departments, highlighting the four most touristic ones.

FIGURE 1.
Average annual international tourists by department (2012-2021)



Source: Colombia's Ministry of Commerce, Industry, and Tourism and authors' calculations.

In terms of economic growth, in 2021, Colombia ranked number 98 on Gross Domestic Product per capita at nominal prices, with a total of $6,644 USD (World Economic, 2022), below the world average and those of its neighboring countries of Brazil and Panama. It ranked similarly to countries of its geographical area such as Peru or Ecuador and higher than certain surrounding countries such as Venezuela.

In terms of the measurement of development, Colombia ranked number 88 on the 2021 edition of the Human Development Index (HDI), with a value of 0.752. It was part of the group of countries having a high HDI (UNDP, 2022), above the global average and that of neighboring countries such as Venezuela. It was similar to other countries in its geographical area such as Brazil, Peru, and Ecuador, and was below neighboring countries such as Panama.

In addition to this data at a national level, Table 1 details the situation of each of the 33 departments in terms of economic growth, economic development, and the importance of tourism for the analyzed period.

Between 2012 and 2021, the departments of Casanare, Guaviare, and Meta had GDP per capita values well above the national average, in contrast to departments such as La Guajira, Vichada, and Vaupes, which had values that were much below the average. Likewise, Bogota, D.C stood out for its high Human Development Index, unlike regions such as Vaupes and Guania. Finally, in terms of tourism, Bogota, D.C and Bolivar stand out when compared to the other departments. All of these data reveal the significant regional disparities at an economic, social, and touristic level in Colombia.

Given the importance of tourist activity in Colombia, which justifies its choice as a case study, as well as the expansion of tourist activity, which has revealed major differences at a regional level, the foundations for this research are evident.

TABLE 1.
Descriptive table of the departments of Colombia (2012-2021 average)

Department	Code	GDP	GDP pc	HDI	Population	Tourism
Amazonas	AMA	606	8.09	0.7017	74,821	4,325
Antioquia	ANT	118,366	18.76	0.7596	6,297,991	268,716
Arauca	ARA	4,501	17.63	0.7310	257,243	390
Atlántico	ATL	35,760	14.34	0.7722	2,492,096	48,225
Bogotá, D.C.	BOG	210,699	28.39	0.8008	7,414,992	898,308
Bolívar	BOL	29,354	14.31	0.7421	2,049,820	307,794
Boyacá	BOY	21,969	18.12	0.7472	1,212,296	4,219
Caldas	CAL	12,754	12.79	0.7656	996,801	7,368
Caquetá	CAQ	3,385	8.40	0.7071	402,746	429
Casanare	CAS	12,882	31.50	0.7379	410,394	690
Cauca	CAU	14,507	10.02	0.7091	1,446,024	3,699
Cesar	CES	14,984	12.83	0.7178	1,171,366	2,339
Chocó	CHO	3,509	6.73	0.6847	523,603	1,879
Córdoba	COR	13,890	7.89	0.7033	1,760,824	2,257
Cundinamarca	CUR	49,143	18.02	0.7599	2,743,450	10,523
Guainía	GUA	304	6.64	0.6630	45,938	91
La Guajira	GJR	672	0.79	0.6844	854,883	3,024
Guaviare	GVR	13,498	167.25	0.7413	81,036	41
Huila	HUI	8,572	7.93	0.7145	1,083,418	3,396
Magdalena	MAG	10,818	8.20	0.7153	1,318,863	21,662
Meta	MET	29,405	29.08	0.7620	1,014,167	3,769
Nariño	NAR	12,246	7.54	0.7107	1,623,306	13,610
Norte de Santander	NSA	12,749	8.63	0.7363	1,478,744	17,602
Putumayo	PUT	3,321	9.85	0.7051	338,861	1,019
Quindío	QUI	6,528	12.18	0.7697	535,534	10,925
Risaralda	RIS	12,945	13.81	0.7494	936,110	21,325
S. Andrés and Providencia	SAD	1,265	20.52	0.7773	61,653	74,526
Santander	SAN	52,422	24.32	0.7622	2,155,367	16,236
Sucre	SUC	6,796	7.59	0.7287	893,852	1,005
Tolima	TOL	17,524	13.17	0.7337	1,330,483	4,960
Valle del Cauca	VCA	80,016	17.95	0.7769	4,453,672	148,175
Vaupés	VAU	237	6.02	0.6371	39,330	32
Vichada	VIC	533	5.09	0.7355	104,812	116

Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

3.2. Data and variables

In this section, the measurement variables are identified and detailed for each of Colombia's 33 departments, with respect to both tourist activity and economic growth and development. Specifically, the following variables are used:

- Tourism. International tourism is used to characterize tourist activity from a demand perspective, measuring the number of foreigners who, in absolute terms, visit each of Colombia's 33 departments, thereby indicating the degree of a region's tourist activity. This information is obtained from Colombia's Ministry of Commerce, Industry, and Tourism.
- Economic growth. The Gross Domestic Product (GDP) is used at constant 2015 prices, measured in millions of pesos, relativized in per capita terms, such that the GDP of each department is divided by the population making up said region. This information is constructed and obtained from the National Administrative Department of Statistics of Colombia (DANE).
- Economic development. The Human Development Index (HDI) is used to characterize the economic development. It is the average result of the values of three dimensions: education, health, and standard of living. Regional data is provided by the Global Data Lab through an adaptation of national-level data provided by the United Nations Development Program.

For the measurement of both economic growth and the population of each of the departments, longer series could have been obtained, beginning the series in a previous year, as with the use of the HDI, which could have included data from the fiscal year of 1990. However, as for the number of international tourists in Colombia at a departmental level, the first data series are available as of 2012, except for some departments which, in an isolated manner, may have some earlier measurements. Furthermore, the HDI data at a regional level are available until 2021. Therefore, the time series used in this work has a duration of 10 years, the period between 2012 and 2021.

4. Methodology

According to the objective of this work, analyzing the relationship between tourism and economic growth and the relationship between tourism and economic development in Colombia at a regional level, two analyses are presented in parallel in this section. On the one hand, there is a qualitative dynamic analysis between tourism and economic growth, and, on the other hand, there is a qualitative dynamic analysis between tourism and economic development.

Clustering algorithms are used to perform this dynamic analysis. When working with tourism, economic growth and economic development series, which have different units of measurement, the commonly used distance measures are not valid. To overcome this problem, symbolic time series analysis was used (Brida et al., 2020). Thus, starting from the concept of economic regime, in a first step, symbolic series are constructed. This allows moving from two-dimensional time series of tourism and economic growth (or tourism and economic development) to one-dimensional time series of tourism and economic growth, which contain all of the relevant information on the dynamics of these variables. Once the symbolic series are obtained for each department of Colombia, a distance is defined. This distance matrix is used as input, to apply a clustering algorithm, which results in homogeneous groups.

4.1. Symbolization

To capture the qualitatively relevant properties, the concepts of regime and regime dynamics are introduced (Brida et al., 2015; Brida et al., 2020). Each regime corresponds to a model of economic performance that differs qualitatively from the others. The partition of the space of the states of tourism and economic growth/economic development is carried out based on the annual values of the indicator used to measure tourism (x) and economic growth/economic development (y).

Thus, the space is divided into four quadrants, determined by the annual values of the indicator used to measure tourism (x) and economic growth/economic development, $\bar{x}_t$ and $\bar{y}_t$ respectively, with $t = 1, \dots, 10$. Based on this division of the space of states in regimes, two types of dynamics are identified: one within each of the regimes and another of change between the regimes. While the dynamics observed in each regime determine a performance model that differs from the models that act in the others, the dynamics of the change from one regime to another indicate where an economy is at each specific moment of time. This dynamic qualitatively describes the performance in terms of tourism and economic growth/economic development.

Regarding the methodological development, in the series symbolization process, the bi-dimensional time series $\{(x_1, y_1),\ (x_2, y_2), \dots, (x_T, y_T)\}$ are substituted by a sequence of symbols: $s = \{s_1, s_2, \dots, s_T\}$, such that $s_t = j$ if, and only if (x_t, y_t) belongs to the R_j region. The regimes are defined as follows:

$$R_1 = \{(x, y) :\ x_t \geq \bar{x}_t\ , y_t \geq \bar{y}_t)\} \tag{1}$$

Regime 1: regions with economic growth/economic development and tourism importance that are higher than the average. In this regime, regions with greater growth or development and with major tourism sector development are expected to be found.

$$R_2 = \{(x, y) : x_t \leq \bar{x}_t\ , y_t \geq \bar{y}_t)\} \tag{2}$$

Regime 2: regions with high economic growth/economic development and low tourism importance. In this regime, it is expected that there will be some regions having a high level of growth or development, but with specialization in other economic activities.

$$R_3 = \{(x, y) : x_t \leq \bar{x}_t\ , y_t \leq \bar{y}_t)\} \tag{3}$$

Regime 3: regions with low economic growth/economic development and low tourism importance. Here, it is expected that regions with lower levels of growth or development will be found, as well as a tourism sector that is not very relevant to its economic activity.

$$R_4 = \{(x, y) : x_t \geq \bar{x}_t\ , y_t \leq \bar{y}_t)\} \tag{4}$$

Regime 4: regions with low economic growth/economic development and high tourism importance. In this region, it is expected that regions with low levels of growth and development and high levels of tourism will be found.

Based on this symbolization, the symbolic distance (the discrete metric of zeros and ones) is defined, and then, based on the constructed distance matrix, a nearest neighbor clustering is made (in this algorithm, the distance between two clusters is defined as the minimum of the element-to-element distances).

4.2. Clustering

In order to detect homogenous groups of regions with similar dynamic behavior in terms of tourism behavior and economic growth/economic development, a cluster of departments was formed. To form this cluster, after obtaining the one-dimensional symbolic series, a metric was introduced to permit comparison of the dynamics of the region and allow for homogenous groups to be obtained. Given two regions, i and j, with symbolic sequences $\{s_{it}\}_{t=1}^{t=T}$ and $\{s_{jt}\}_{t=1}^{t=T}$ the distance between them is defined as follows:

$$d(s_i, s_j) = \sum_{t=1}^{T} f(s_{it}\ , s_{jt}), \qquad where\ f(s_i\ , s_j) = \begin{cases} 0\ if\ s_{it} = s_{jt} \\ 1\ if\ s_{it} \neq s_{jt} \end{cases} with\ t = 1, \dots, T \tag{5}$$

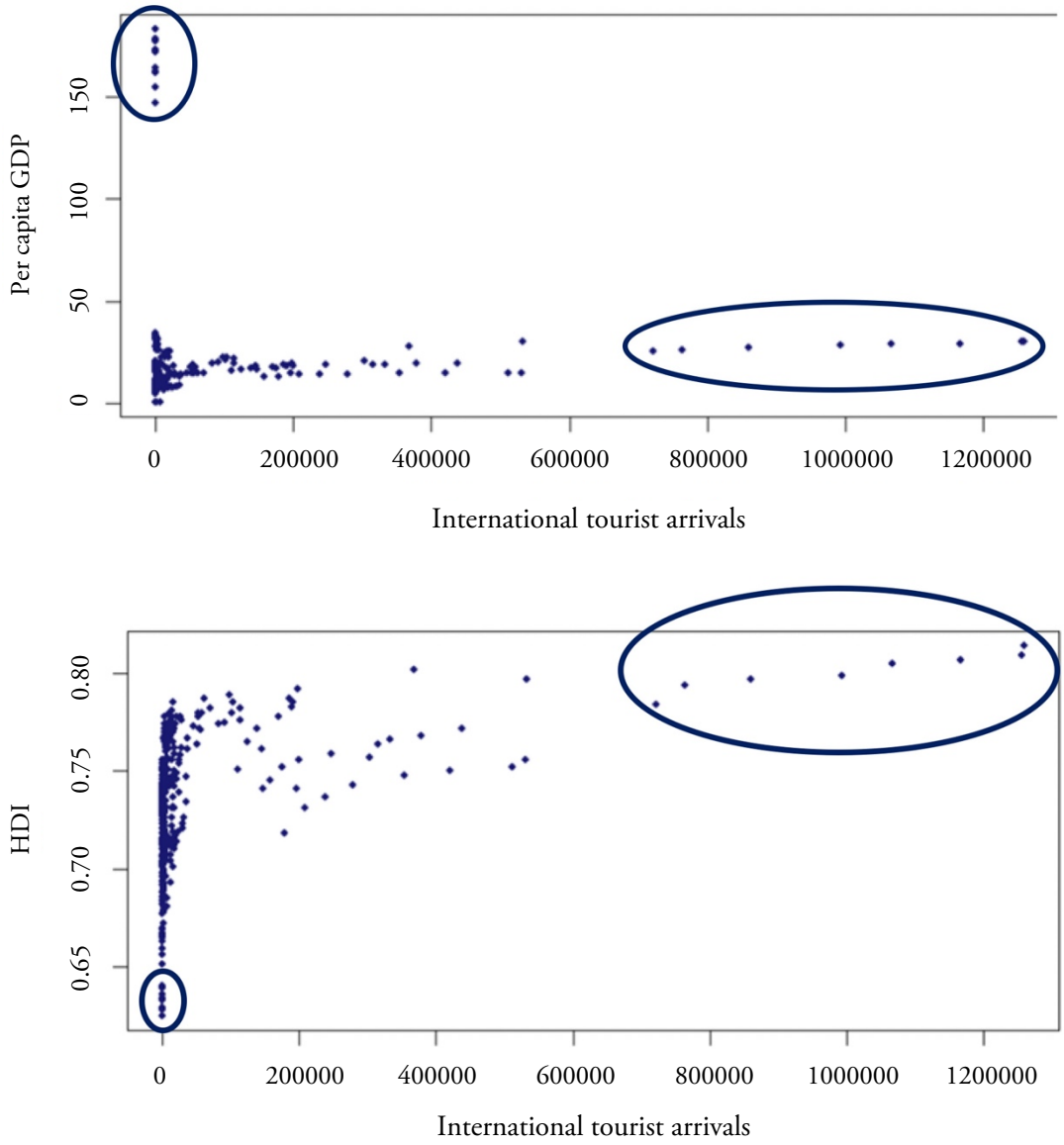
Once the distance matrices are obtained in each case, the nearest neighbor method is used to form the clusters (Mantegna & Stanley, 1999).

5. Empirical results

5.1. Symbolic time series analysis

Figure 2 shows the point cloud corresponding to the 2012-2021 period, with the respective values of tourism and economic growth (upper panel) and tourism and economic development (lower panel). Each point represents a department for each of the years of the time series (in the first case, tourism and economic growth, and in the second case, tourism and economic development).

Figure 2.
Point cloud for the 33 departments (2012-2021)



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

In the first part of the figure (tourism and economic growth), there is a large concentration at low levels of both international tourism and economic growth. The points having high GDP per capita

correspond to Guaviare, the only department with a GDP per inhabitant exceeding 150 million Colombian pesos per year. On the other hand, the points having more than 600,000 international tourist arrivals correspond to Bogota, which is the department receiving the most international tourism in Colombia.

In the second part of the figure (tourism and economic development), there is a large concentration around the average level of development in Colombia. The points having a low level of HDI correspond to Vaupes, the only department with an HDI under 0.65. On the other hand, the points with over 600,000 international tourists received correspond to Bogota, as mentioned previously.

Table number 2 shows the percentage of permanence in each regime for each department during the analyzed period, with the left side presenting the relationship between tourism and economic growth and the right side showing that of tourism and economic development.

TABLE 2.
Proportion of permanence of the departments in each regime

Department	Tourism and economic growth				Tourism and economic development			
	R1	R2	R3	R4	R1	R2	R3	R4
AMA	0%	0%	**100%**	0%	0%	0%	**100%**	0%
ANT	70%	0%	0%	30%	**100%**	0%	0%	0%
ARA	0%	30%	70%	0%	0%	30%	70%	0%
ATL	0%	0%	80%	20%	20%	80%	0%	0%
BOG	**100%**	0%	0%	0%	**100%**	0%	0%	0%
BOL	0%	0%	0%	**100%**	**100%**	0%	0%	0%
BOY	0%	60%	40%	0%	0%	**100%**	0%	0%
CAL	0%	0%	**100%**	0%	0%	**100%**	0%	0%
CAQ	0%	0%	**100%**	0%	0%	0%	**100%**	0%
CAS	0%	**100%**	0%	0%	0%	**100%**	0%	0%
CAU	0%	0%	**100%**	0%	0%	0%	**100%**	0%
CES	0%	0%	**100%**	0%	0%	0%	**100%**	0%
CHO	0%	0%	**100%**	0%	0%	0%	**100%**	0%
COR	0%	0%	**100%**	0%	0%	0%	**100%**	0%
CUR	0%	50%	50%	0%	0%	**100%**	0%	0%
GUA	0%	0%	**100%**	0%	0%	0%	**100%**	0%
GJR	0%	0%	**100%**	0%	0%	0%	**100%**	0%
GVR	0%	**100%**	0%	0%	0%	**100%**	0%	0%
HUI	0%	0%	**100%**	0%	0%	0%	**100%**	0%
MAG	0%	0%	**100%**	0%	0%	0%	**100%**	0%
MET	0%	**100%**	0%	0%	0%	**100%**	0%	0%
NAR	0%	0%	**100%**	0%	0%	0%	**100%**	0%
NSA	0%	0%	**100%**	0%	0%	**100%**	0%	0%
PUT	0%	0%	**100%**	0%	0%	0%	**100%**	0%
QUI	0%	0%	**100%**	0%	0%	**100%**	0%	0%
RIS	0%	0%	**100%**	0%	0%	**100%**	0%	0%
SAD	70%	20%	0%	10%	80%	20%	0%	0%

TABLE 2. CONT.
Proportion of permanence of the departments in each regime

Department	Tourism and economic growth				Tourism and economic development			
	R1	R2	R3	R4	R1	R2	R3	R4
SAN	0%	**100%**	0%	0%	0%	**100%**	0%	0%
SUC	0%	0%	**100%**	0%	0%	10%	90%	0%
TOL	0%	0%	**100%**	0%	0%	90%	10%	0%
VCA	50%	0%	0%	50%	**100%**	0%	0%	0%
VAU	0%	0%	**100%**	0%	0%	0%	**100%**	0%
VIC	0%	0%	**100%**	0%	0%	70%	30%	0%

Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

As seen in the table, the large majority of the departments remain in the same regime throughout the entire period: 26 departments for the case of tourism and growth, 27 departments for the case of tourism and development. Therefore, changes in regime are difficult to find since these changes tend to imply a structural alteration in the economy. Moreover, in a short period of time (such as the period analyzed in this work, which spans 10 years), it would be unusual to encounter this type of structural change.

Furthermore, in both cases, regime 3 (departments with low tourism importance and low levels of growth and development) is the one in which the most departments predominate: 20 departments for the case of tourism and growth, 10 departments for the case of tourism and development. This suggests that, although tourism as an economic activity is important for Colombia, only certain regions of the country specialize in it and use its expansion to channel greater economic growth and improved host population development.

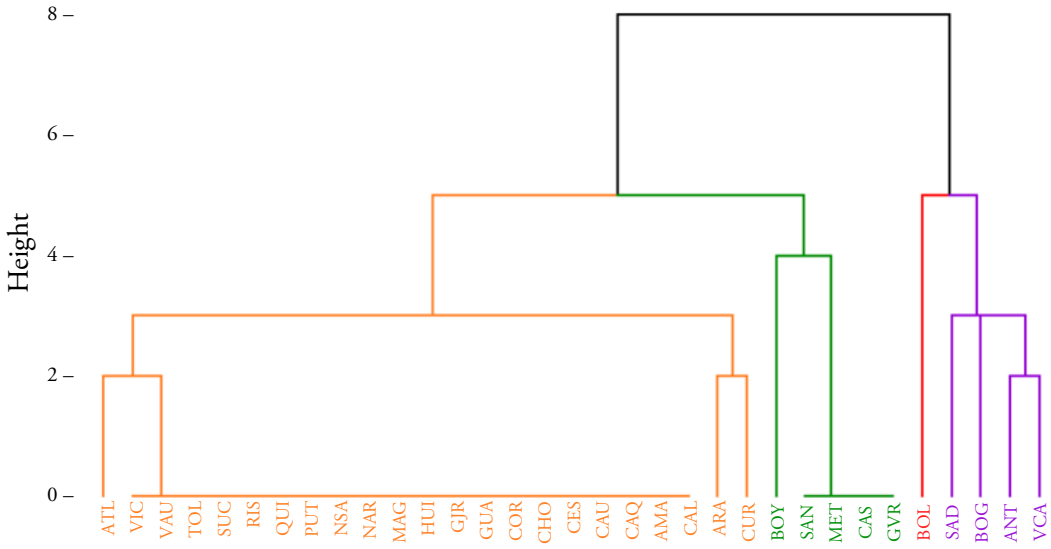
5.2. FORMATION OF CLUSTERS BASED ON TOURISM BEHAVIOR AND REGIONAL GROWTH

5.2.1. TOURISM AND ECONOMIC GROWTH

Firstly, a dendrogram is shown for the grouping of departments displaying similar dynamic behavior in terms of tourism, measured through the number of international tourists, and economic growth, based on the GDP per capita (Figure 3).

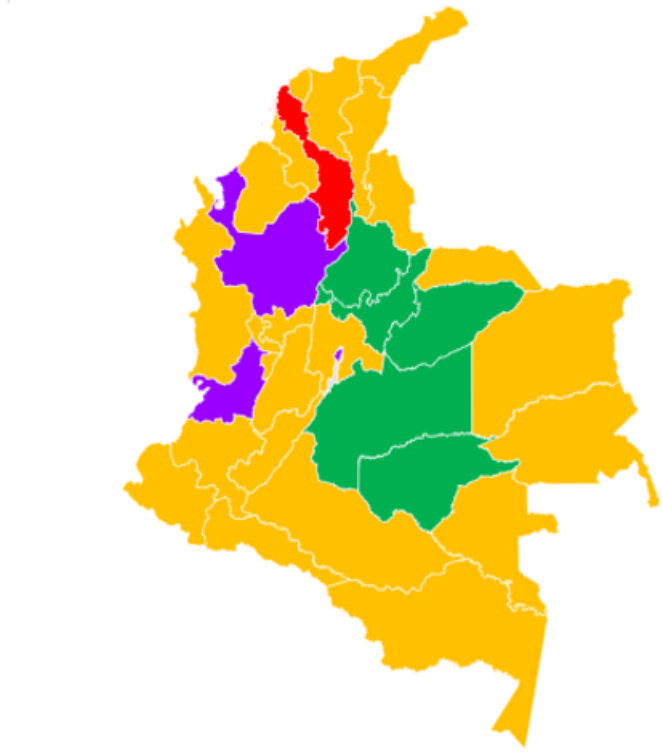
From this grouping process, 3 groups of regions are obtained, which share homogenous dynamic behavior in terms of tourism activity. Figure 4 shows the clustered regions sharing a similar behavior and having significant differences with respect to the other homogenous groups.

FIGURE 3.
Dendrogram of clustering of departments between tourism and economic growth



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

FIGURE 4.
Homogenous groups of departments with similar dynamic behavior (tourism – economic growth)



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations. (Orange: group A / Purple: group B / Green: group C / Red: outliers)

As shown in the previous figure, three main groups exist. They are characterized by the performance of these variables (tourism and economic growth). Bolivar is the only department that is not grouped, that is, that does not connect with any of the other main groups.

Table 3 below summarizes what is shown in the previous map.

Table 3.
Homogenous groups of departments with similar dynamic behavior (tourism – economic growth)

Group A (Orange)	Group B (Purple)	Group C (Green)
Departments with low tourism activity and low GDP levels	Departments with high tourism activity and a high GDP level	Departments with low tourism activity and high GDP levels
23 departments	4 departments	5 departments
Amazonas, Arauca, Atlántico, Caldas, Caquetá, Cauca, Cesar, Chocó, Córdoba, Cundinamarca, Guainía, La Guajira, Huila, Magdalena, Nariño, Norte de Santander, Putumayo, Quindío, Risaralda, Sucre, Tolima, Vaupés and Vichada	Antioquia, Bogotá, San Andrés and Providencia and Valle del Cauca	Boyacá, Casanare, Guaviare, Meta and Santander
These regions are characterized by low numbers of tourists and by not being specialized in other productive activities that would increase their GDP level.	These regions receive a large number of tourists, resulting in their tourism specialization. Thus, high levels of economic growth are generated from the intense economic activity taking place in the regions.	Regions specialize in other productive activities that allow them to attain high levels of economic growth.

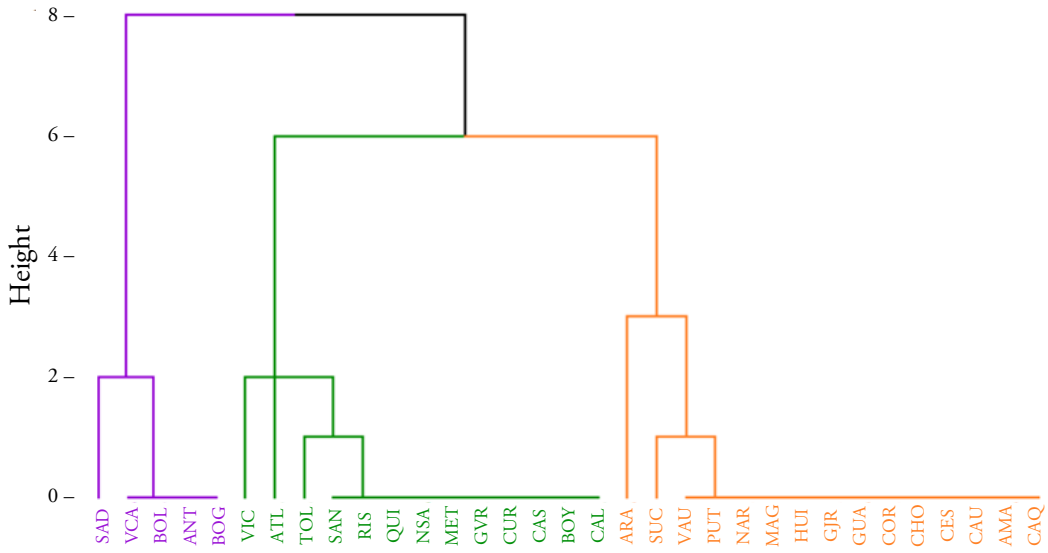
Source: Author's own creation.

5.2.2. Tourism and economic development

Second, a dendrogram shows the grouping of departments displaying similar dynamic behavior in terms of tourism, measured based on number of international tourists, and economic development, measured based on the HDI (Figure 5).

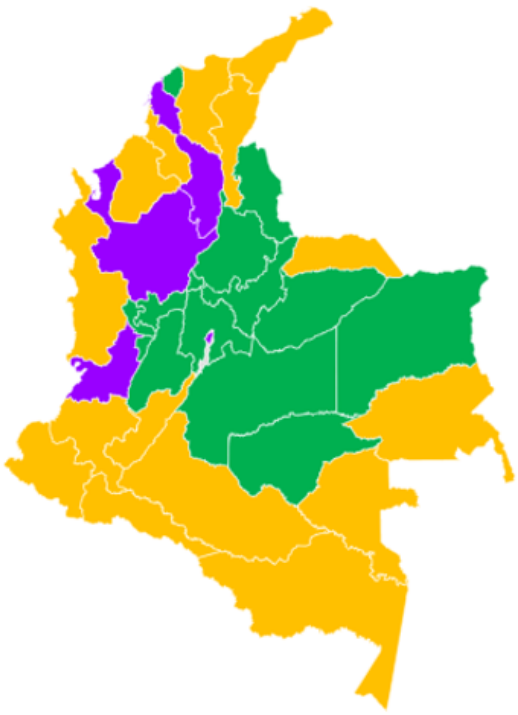
From this grouping process, 3 groups of regions sharing homogenous dynamic behavior in terms of tourism activity are obtained. Figure 6 presents the grouped regions displaying similar behavior and significant differences with respect to the other homogenous groups.

FIGURE 5.
Dendrogram of the grouping of departments between tourism and economic development



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

FIGURE 6.
Homogenous groups of departments displaying similar dynamic behavior (tourism – economic development)



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations. (Orange: group A / Purple: group B / Green: group C).

As seen in the previous figure, three main groups exist, characterized by their performance on these variables (tourism and economic development).

Table 4.
Homogenous groups of departments with similar dynamic behavior (tourism – economic development)

Group A (Orange)	Group B (Purple)	Group C (Green)
Departments with low tourism activity and low GDP levels	Departments with high tourism activity and a high GDP level	Departments with low tourism activity and high GDP levels
15 departments	5 departments	13 departments
Amazonas, Arauca, Caquetá, Cauca, Cesar, Chocó, Córdoba, Guainía, La Guajira, Huila, Magdalena, Nariño, Putumayo, Sucre and Vaupés	Antioquia, Bogotá, Bolívar, San Andrés and Providencia and Valle del Cauca	Atlántico, Boyacá, Caldas, Casanare, Cundinamarca, Guaviare, Meta, Norte de Santander, Quindío, Risaralda, Santander, Tolima and Vichada
These departments are experiencing a developmental trap, given their low levels of development, which hinder the expansion of tourism activity. Simultaneously, this lack of tourism development hinders the improvement of the host population's living conditions.	The tourism specialization of these departments serves as a tool for economic development.	Although tourism is not a fundamental sector of the economy, other economic activities allow the departments to attain high levels of development. .

Source: Author's own creation.

5.2.3. Regional tourism characterization as a determinant of economic growth and development

When considering both international tourism and economic growth, as well as international tourism and economic development, 3 groups of departments displaying similar behavior are obtained. However, when altering between economic growth and economic development, their composition differs.

On the one hand, the relationships between tourism and economic growth, as well as the relationships between tourism and economic development, appear to behave quite similarly, given that, of the 33 departments analyzed, group changes only take place in 9 of these (and all in the same direction):

- 8 regions that, when analyzing tourism and economic growth, are part of Group A (low tourism activity and low economic growth levels), but when analyzing tourism and economic development, are part of Group C (low level of tourism activity and high level of development). These departments are: Atlántico, Caldas, Cundinamarca, Norte de Santander, Quindio, Risaralda, Tolima, and Vichada.
- 1 region does not fit into any of the groups when analyzing tourism and economic growth (outlier), but when analyzing tourism and economic development is part of group B (high level of tourism activity and high level of development). This department is Bolivar.

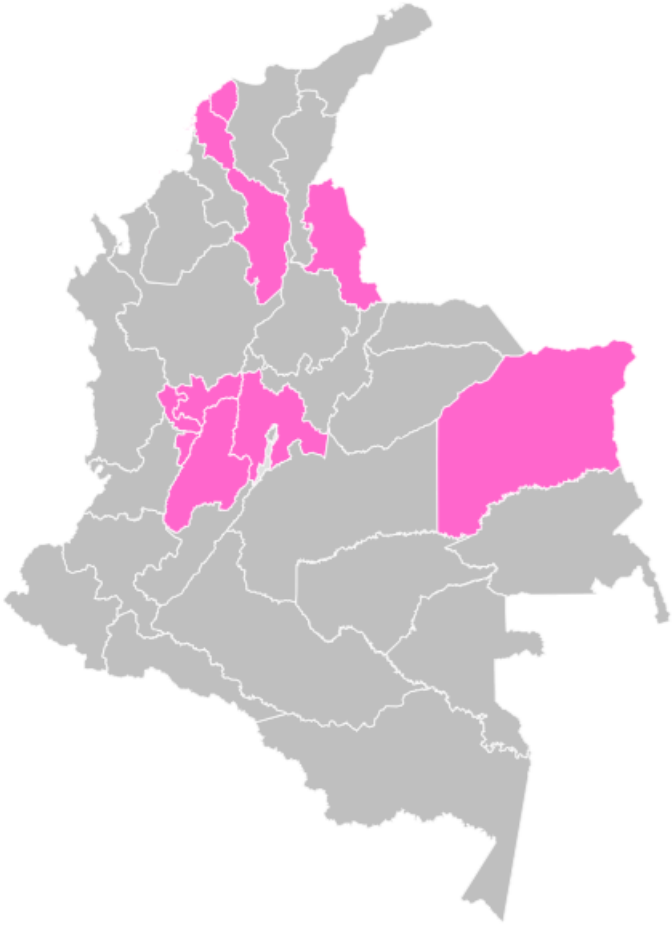
First, it should be noted that a higher level of tourism activity means a higher level of economic growth for the departments, and, most importantly, this economic growth generated by tourism also results in improved living conditions in the resident population and, therefore, a higher level of economic development.

Second, it should also be highlighted that no departments which, having increased their levels of tourism activity, do not display high levels of economic growth and high levels of development. Therefore,

it should be noted that tourism activity is considered a tool of growth and development at a regional level in Colombia.

It would be interesting to characterize these departments in comparison with the other regions, in order to identify the tourism factors that are determinant to achieving higher levels of economic growth and development at a regional level in Colombia (Table 5).

FIGURE 7.
Groups of regions with similar behavior according to economic growth or development



Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.
(Pink: regions that change regime when analyzing tourism-growth or tourism-development / Gray: regions with no change).

Regarding tourist characterization based on demand data, tourist activity level is relevant for those departments in which tourism represents a higher level of economic growth and development, both in absolute terms (with a difference of 100,000 annual tourists in average terms between the lower limit of the Group B departments and the other departments) as well as in relative terms (a difference of 0.02 tourists per capita between both limits).

As for tourist characterization according to supply data, the number of travel agencies is the only indicator that is relevant for departments where tourism represents higher levels of growth and economic development (with a difference of an average of 155 travel agencies between the lower limit of the Group B departments and the other departments), since both in availability of accommodation (measured by the number of tourist lodging establishments and the number of tourism properties) and dining establishments, departments in Group B do not display higher figures than the other departments.

TABLE 5.
Characterization of regions in which tourism is related to higher levels of economic growth and higher levels of development

B Gr. (high tourism/high growth-development)	Other regions
Bogota, Bolivar(*), Antioquia, Valle del Cauca and San Andres and Providencia (1)	
International tourism: > 148.000 tourists-year	International tourism: ≤ 48.000 tourists-year
Tourist intensity: > 0.05 tourists pc	Tourist intensity: ≤ 0.03 tourists pc
Travel agencies: ≥ 680 annual mean(1)	Travel agencies: < 525 annual mean
Lodging establishments: > 465 annual mean	Lodging establishments: < 1,075 annual mean
Tourism properties: > 270 annual mean	Tourism properties: < 910 annual mean
Dining establishments: > 70 annual mean	Dining establishments: < 95 annual mean

(*) Although considered an outlier when analyzing tourism and economic growth, it could be grouped at distance 5 with Group B. (1) Although it is outside of the absolute values indicated in some relative indicators (per capita) it has the highest values on those indicators (for example, in tourist intensity, the only one with a value > 1; specifically, 1.21).
Source: Colombia's Ministry of Commerce, Industry, and Tourism; National Administrative Department; United Nations Development Program and authors' calculations.

6. CONCLUSION

This work advances the study of the interaction between tourism, economic growth, and economic development, taking an innovative approach that addresses two fundamental aspects: whether the economic growth generated by tourist activity represents an improvement in development and following a regional-level approach for this type of research. To achieve the objectives, qualitative dynamic analyses have been performed, introducing the concept of regime and the dynamics of regimes in order to differentiate between the qualitatively relevant properties of each of them. Subsequently, a grouping of departments was carried out to detect homogeneous groups of regions displaying similar dynamic behavior in terms of tourist behavior and economic growth/economic development. The techniques used allowed us to understand dynamic and complex processes that may go unnoticed in static analyses, providing guidance for future policies and practices in the field of tourism and regional development.

The relationship between tourism and the dynamics of economic growth and development in Colombia reveals major differences at a regional level. This study has enabled the identification of homogeneous groups of regions displaying similar behavior in terms of tourism and economic growth and in terms of tourism and economic development. Few differences have been observed in both classifications, with minimal changes occurring in the same direction (for example, departments that, despite not being specialized in tourism, present low growth levels but high development levels).

It should be noted that, in all cases, those departments having the highest levels of tourism (Bogota, Bolivar, Antioquia, Valle del Cauca, and San Andres and Providencia) display a high level of economic growth as well as a high level of economic development. Relevantly, there are no exceptions to this result (that is, no departments had elevated levels of tourism and low levels of economic growth or low development levels), suggesting that the expansion of tourism activity may enable economic growth and development at a regional level in Colombia.

The characterization of the Colombian departments made from a tourism perspective reveals that factors related to tourism demand (tourism activity and intensity) are relevant in the Group B regions (those with high tourism and high levels of economic growth and development), whereas the greater provision of factors related to the tourism offer is almost irrelevant.

As for policy recommendations, two fundamental aspects should be highlighted. On the one hand, it is necessary to develop policies related to attracting international tourists to the departments that are less specialized in this activity, since it has been shown that the most touristy regions have high levels of

economic growth and development. On the other hand, policies should be developed to provide resources or attraction factors for tourists, as opposed to policies attempting to increase the complementary tourism offer, given that this is not relevant to the relationship between tourism and economic growth-development.

Limitations of this work include the short time period that was analyzed, due to data availability, and the absence of additional indicators to characterize tourism demand at a regional level: motivation, stay, and tourist expenditure, among others.

Future lines of research could include an econometric analysis to determine the causal relationships existing between the variables examined in this work using qualitative dynamic analysis. It may also be possible to expand upon the characterization of the departments based on other indicators that are not exclusively touristic in nature.

References

Alcalá-Ordóñez, A., Brida, J. G., & Cárdenas-García, P. J. (2023). Has the tourism-led growth hypothesis been confirmed? Evidence from an updated literature review. *Current Issues in Tourism*, 1-37. https://doi.org/10.1080/13683500.2023.2272730

Alcalá-Ordóñez, A., & Segarra, V. (2023). Tourism and economic development: A literature review to highlight main empirical findings. *Tourism Economics*, *31*(1). https://doi.org/10.1177/13548166231219638

Balaguer, J., & Cantavella-Jordá, M. (2002). Tourism as a long-run economic growth factor: the Spanish case. *Applied Economics*, *34*(7), 877-884. https://doi.org/10.1080/00036840110058923

Bonet, J., & Meisel, A. (2009). Regional economic disparities in Colombia. *Journal of Regional Research*, *14*, 61-80.

Bassil, C., Harb, G., & Al Daia, R. (2023). The economic impact of tourism at regional level: A systematic literature review. *Tourism Review International*, *27*(2), 159-175.

Brida, J. G., Brindis, M. A. R., & Mejía-Alzate, M. L. (2020). La contribución del turismo al crecimiento económico de la ciudad de Medellín - Colombia. *Revista de Economia Del Rosario*, *24*(1), 1-23. https://doi.org/10.12804/revistas.urosario.edu.co/economia/a.8926

Brida, J. G., Garrido, N., & Matesanz-Gómeza, D. (2015). Análisis jerárquico de la dinámica económica de las comunidades españolas en el periodo 1955-2009. *Investigaciones Regionales - Journal of Regional Research*, *31*, 121-144.

Brida, J. G., & Monterubbianesi, P. D. (2010). Causality Between Economic Growth and Tourism Expansion: Empirical Evidence from Some Colombian Regions. *Journal of Tourism Challenges and Trends*, *3*(1), 153-164.

Brida, J. G., Monterubbianesi, P. D., Zapata-Aguirre, S., & Giraldo Velásquez, C. M. (2010). Turismo y Crecimiento Económico: El Caso de Antioquia. *Turismo y Desarrollo Local*, *7*. https://EconPapers.repec.org/RePEc:erv:turdes:y:2010:i:7:11

Brida, J. G., Pereyra, J. S., Risso, W. A., Such Devesa, M. J., & Zapata Aguirre, S. (2009). The Tourism-led Growth Hypothesis: Empirical Evidence from Colombia. *Tourismos: An International Multidisciplinary Journal of Tourism*, *4*(2), 13-27.

Brida, J. G., Cortes-Jimenez, I., & Pulina, M. (2016). Has the tourism-led growth hypothesis been validated? A literature review. *Current Issues in Tourism*, *19*(5), 394-430. https://doi.org/10.1080/13683500.2013.868414

Brida, J. G., Rodríguez Brindis, Mejía Alzate, M. L., & Zapata Aguirre, S. (2017). La contribución directa del turismo al crecimiento económico de Colombia: Análisis por ramas características del sector utilizando la Cuenta Satélite de Turismo-CST. *Revista de Estudios Regionales*, *109*, 121-138.

Brida, J.G., Matesanz, D., & Segarra, V. (2020). On the empirical relationship between tourism and economic growth. *Tourism Management*, *81*. https://doi.org/10.1016/j.tourman.2020.104131

Calero, C., and Turner, L.W. (2020). Regional economic development and tourism: A literature review to highlight future directions for regional tourism research. *Tourism Economics*, *26*(1), 3-26. https://doi.org/10.1177/1354816619881244

Cárdenas-García, P. J., Sánchez-Rivero, M., & Pulido-Fernández, J. I. (2015). Does tourism growth influence economic development? *Journal of Travel Research*, *54*(2), 206-221. https://doi.org/10.1177/0047287513514297

Cárdenas-García, P. J., & Pulido-Fernández, J. I. (2019). Tourism as an economic development tool. Key factors. *Current Issues in Tourism*, *22*(17), 2082-2108. https://doi.org/10.1080/13683500.2017.1420042

Cotte Poveda, A. (2011). Economic development and growth in Colombia: An empirical analysis with super-efficiency DEA and panel data models. *Socio-Economic Planning Sciences*, *45*(4), 154-164. https://doi.org/10.1016/j.seps.2011.07.003

Croes, R., Ridderstaat, J., Bak, M., and Zientara, P. (2021). Tourism specialization, economic growth, human development and transition economies: The case of Poland. *Tourism Management*, *82*. https://doi.org/10.1016/j.tourman.2020.104181

Delgado Munévar, W. G. (2023). Turismo y crecimiento económico mediante cointegración de datos panel para países latinoamericanos. *Turismo y Sociedad*, *33*, 73-83. https://doi.org/10.18601/01207555.n33.03

Galvis-Aponte, L. A., Galvis-Larios, W., & Hahn-De-Castro, L. W. (2017). *Una Revisión de los Estudios de Convergencia Regional en Colombia*.

Garavito, L., & Ochoa, F. (eds.) (2016). *Debates contemporáneos sobre turísmo, tomo I. Nuevos retos del turismo. Casos de estudio sobre vínculos entre la gobernanza, el desarrollo y la sostenibilidad*. (1st ed.). Universidad del Externado. https://doi.org/10.2307/j.ctv18mspqw

Garavito, L., Rozo, E., & Mojica, F. (2019). ¿Puede el turismo liderar el desarrollo competitivo y sostenible en el presente y el futuro próximo en Colombia? *Via Tourism Review*, *15*. https://doi.org/10.4000/viatourism.3522

Lee, Y. S. (2017). General theory of law and development. *Cornell International Law Review*, *50*(3), 432-435. https://scholarship.law.cornell.edu/cilj/vol50/iss3/2

Meng-Yi Tai, S.-W. H., & Chao, C.-C. (2023). Pandemic, tourism, and the economy: a tourism-revenue target zone approach. *Asia-Pacific Journal of Accounting & Economics*, *30*(6), 1659-1674. https://doi.org/10.1080/16081625.2023.2186902

Oficina de Estudios Económicos. (2024). *Informe Mensual de Turismo Noviembre 2023*.

Peiró-Palomino, J., Picazo-Tadeo, A. J., & Tortosa-Ausina, E. (2021). Measuring well-being in Colombian departments. The role of geography and demography. *Socio-Economic Planning Sciences*, *78*. https://doi.org/10.1016/j.seps.2021.101072

Risso, W. A. (2018). Tourism and economic growth: A worldwide study. *Tourism Analysis*, *23*(1), 123-135.

Rodríguez-Benavides, D., & Ceballos-Mina, O. (2022). Clubes de convergencia regional en Colombia 2000-2016: un análisis flexible por departamentos. *Investigaciones Regionales - Journal of Regional Research*, *53*, 47-65.

Rodríguez-Victoria, O. E., González-Loureiro, M., & Puig, F. (2017). Economic Competitiveness: Effects of Clustering, Innovation Strategy and the Moderating Role of Location in the Colombian Hotel Industry. *Investigaciones Regionales - Journal of Regional Research*, *39*, 81-97.

Rozo Bellón, E., & Garavito González, L. (2014). Tourism in Colombia. In *Tourism as an Instrument for Development: A Theoretical and Practical Study (Bridging Tourism Theory and Practice)* (Vol. 5, pp.

211-226). Emerald Group Publishing Limited. https://doi.org/10.1108/S2042-144320140000005027

Sanabria-Gómez, S.A. (2017). Progreso tecnológico y divergencias regionales: evidencia para Colombia (1980-2010). *Investigaciones Regionales - Journal of Regional Research*, *38*, 7-25.

Sanguinet, E., Galvis-Aponte, L. A., Araújo, I. F., & Haddad, E. A. (2020). *Viajeros urbanos de paseo por la playa: efectos regionales del turismo interno en Colombia.*

Such-Devesa, M. J., Zapata-Aguirre, S., Risso, W. A., Brida, J. G., & Pereyra, J. S. (2009). Turismo y Crecimiento Economico: Un Analisis Empirico de Colombia. *Estudios y Perspectivas En Turismo*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1092407

Todaro, M. P., and Smith, S. C. (2020). *Economic Development* (13th Edition). Addison Wesley.

UNDP (2022). *Human Development Index (HDI). UNDP Data Center.* https://hdr.undp.org/data-center/human-development-index#/indicies/HDI

UNWTO United Nations World Tourism Organization (2018). *Tourism and the Sustainable Development Goals: Good Practices in the Implementation of the 2030 Agenda*.

World Bank (2019). *The World Bank Group's Support to Tourism for Sustainable Development.* World Bank.

Wong, D. W. H., Tai, A. C. L., Chan, D. Y. T., & Lee, H. F. (2023). Can tourism development and economic growth mutually reinforce in small countries? Evidence from Singapore. *Current Issues in Tourism*, *27*(8), 1316-1331. https://doi.org/10.1080/13683500.2023.2213879

World Travel & Tourism Council. (2023). *Travel & Tourism Economic Impact 2023-Colombia*.

Zapata-Aguirre, S., López-Zapata, L., & Mejía-Alzate, M. L. (2020). Tourism development in Colombia: between conflict and peace. In C. Monterrubio, K. Andriotis, & D. Stylidis (Eds.), *Tourism planning and development in Latin America* (pp. 49-65). CAB International. https://doi.org/10.1079/9781789243048.0049

ORCID

Juan Gabriel Brida https://orcid.org/0000-0002-2319-5790

Pablo Juan Cárdenas-García https://orcid.org/0000-0002-1779-392X

María Leivy Mejía-Alzate https://orcid.org/0000-0003-4631-8113

Verónica Segarra https://orcid.org/0000-0003-0436-3303

Investigaciones Regionales – Journal of Regional Research (2026/1) 64: 197-213
https://doi.org/10.38191/iirr-jorr.25.004

Articles

Country's governance quality impact on cloud computing adoption in the EU enterprises

*Lejla Karamujic**

Received: 29 May 2024
Accepted: 24 March 2025

Abstract:

Cloud computing is widely recognised as a key technology that enables digitalisation, fuelling the digital economy growth. Using World Bank Governance Indicators, in this paper we examine how country's quality of governance (or good governance) impacts the use of cloud computing in the enterprises in the EU, hence affects digitalisation of the EU. In an empirical examination of 27 EU countries over a period from 2014 to 2021, it was found that good governance positively affects cloud computing adoption in the EU enterprises. In support of the institutional theory, we show that the improving country's governance quality is an important step in accelerating cloud computing adoption by the EU enterprises, the finding providing an important input to the EU government in their execution of the Industrial Strategy 2030, in which digitalisation and cloud computing adoption play an important role. Businesses operating in the EU countries with higher quality of governance are more likely to implement cloud computing in their operations, a cost and performance benefiting technology, a valuable insight to the EU enterprises, which gain direct benefits from increased cloud computing adoption in their businesses.

Keywords: Institutions; cloud computing adoption; digitalisation; good governance; country's governance quality; EU (Europe); World Bank governance indicators; quality of government index.
JEL Classification: O32; O43; O52.

El impacto de la calidad de la gobernanza de un país en la adopción de la computación en la nube en las empresas de la UE

Resumen:

La computación en la nube es ampliamente reconocida como una tecnología clave que facilita la digitalización e impulsa el crecimiento de la economía digital. Utilizando los Indicadores de Gobernanza del Banco Mundial, en este documento examinamos cómo la calidad de la gobernanza (o buena gobernanza) de un país impacta en el uso de la computación en la nube en las empresas de la UE y, por consiguiente, en su digitalización. En un análisis empírico de 27 países de la UE entre 2014 y 2021, se concluyó que la buena gobernanza influye positivamente en la adopción de la computación en la nube en las empresas de la UE. En apoyo de la teoría institucional, demostramos que la mejora de la calidad de la gobernanza de un país es un paso importante para acelerar la adopción de la computación en la nube por parte de las empresas de la UE. Este hallazgo proporciona una importante información al gobierno de la UE en la ejecución de la Estrategia Industrial 2030, en la que la digitalización y la adopción de la computación en la nube desempeñan un papel fundamental. Las empresas que operan en países de la UE con una gobernanza de mayor calidad tienen mayor probabilidad de implementar la computación en la nube en sus operaciones, una tecnología que beneficia tanto a los costes como al rendimiento, lo que constituye una valiosa perspectiva para las empresas de la UE, que se benefician directamente de una mayor

* PhD Student. University of Vienna. Austria. a11913441@unet.univie.ac.at / andrew.lejla@hotmail.com
Corresponding author: a11913441@unet.univie.ac.at / andrew.lejla@hotmail.com

adopción de la computación en la nube en sus negocios.

Palabras clave: Instituciones; adopción de la computación en la nube; digitalización; buena gobernanza; calidad de la gobernanza del país; UE (Europa); indicadores de gobernanza del Banco Mundial; índice de calidad de gobierno.

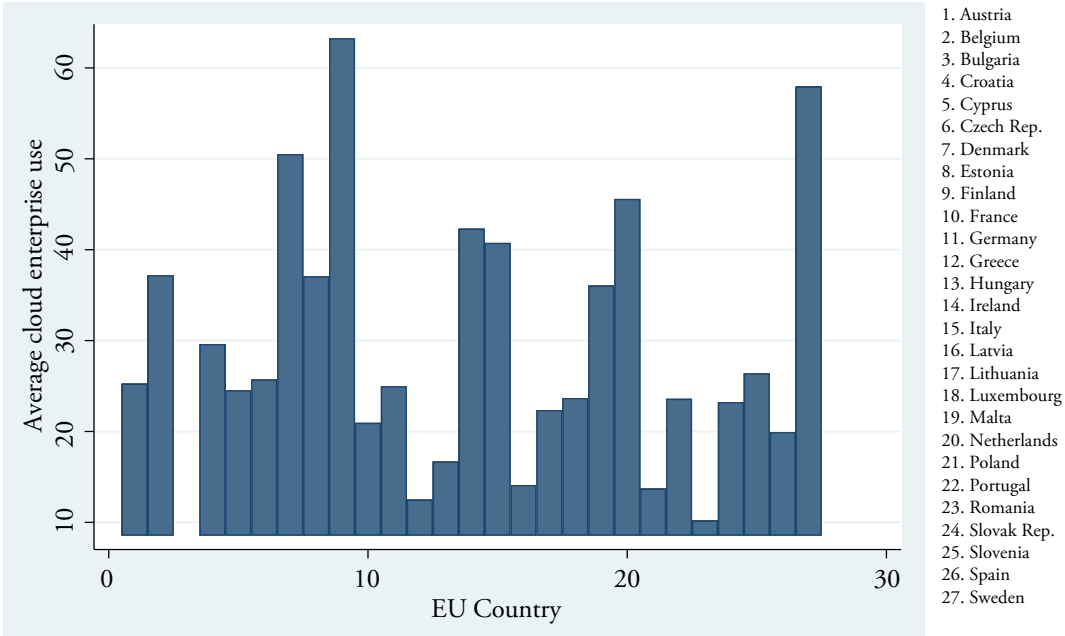
Clasificación JEL: O32; O43; O52.

1. Introduction

Digitalisation is extremely important to growth and development of countries; hence EU has been promoting the development of digital economy. In 2015, the EU Commission launched the Single Digital Market Strategy to assess the maturity of processes related to digitalisation, and in 2021, the Digital Europe Program was launched, with the aim to financially support the digitalisation of the EU countries. Cloud computing[1] is widely recognised as a key technology that enables digitalisation in our economies and societies. The business logic of using cloud computing resides on the premise that organisations can outsource their IT tasks and services at lower cost and higher efficiency. 'Cloud computing has some advantages for individual users; however, it has greater benefits for businesses/enterprises' (Karamujic, 2025). Adoption of cloud computing by enterprises is one of the main objectives of the EU Commission Industrial Strategy 2030 with the aim of accelerating the digitalisation and transforming the EU into 'the continent with a high share of digitalised business.' The specific goal that the Commission set is for 75% of the EU enterprises to have taken cloud computing services[2] by 2030.

Eurostat records the data on the use of cloud computing service in enterprises in the EU, with the Figure 1 below depicting the average use of cloud computing services in enterprises in each of the EU countries in the period 2014 to 2021.

Figure 1.
Average use of cloud computing services in enterprises in the EU

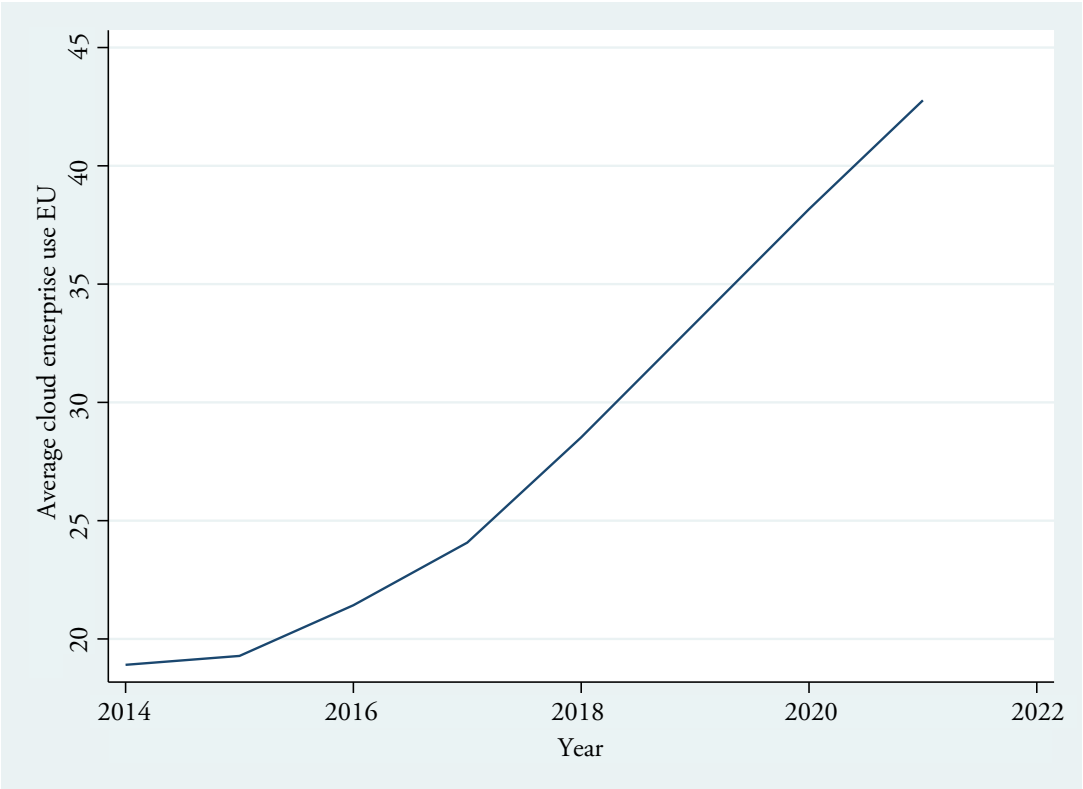


[1] 'Cloud computing is captured by the characteristics of on-demand computing and storage, with shared resources between multiple customers' (Vithayathil, 2017, p.1). 'Cloud computing enables the capable and deliverable IT as services to external customers using internet technologies' (Plummer, D.C. et al., 2008).

[2] Cloud computing is one of the key technologies enabling digital transformation. The other digital technologies mentioned in the Commission Industrial Strategy 2030 are Big Data and AI (Artificial Intelligence).

As it can be observed in the Figure 1 above, all the EU countries are still well below the goal of 75% use of cloud computing by enterprises to be reached by 2030, with 2021[3] average across the union being just above 40% (see Graph 1 below (Source: Karamujic, 2025, 'Does Gaia-X foster cloud computing adoption by European (EU) businesses?').

GRAPH 1.
Average % of EU enterprises using cloud computing services



Additionally, in Figure 1 above, one observes large deviations in the levels of adoption of cloud computing across the EU countries. For example, countries such as Finland (63%), Sweden (58%) and Denmark (51%) have highest levels of adoption of cloud computing services in their enterprises, while countries such as Bulgaria (9%), Romania (11%) and Greece (13%) have some of the lowest levels of the adoption in the EU.

Cloud computing has many benefits to businesses/firms/enterprises (Karamujic, 2025; Chen, Chaung & Nakatani, 2016; Xue & Xin, 2016), in terms of cost and performance benefits, and as such is referred to as business-oriented technology. Considering cloud computing benefits to businesses in general, regardless of national boundaries, one would not expect to see such a large deviation across the EU countries when it comes to the use of cloud computing services.

The EU, which is currently comprised of 27 independent countries, is governed under the set of common rules and regulations under the EU governance, with many laws and policies in the EU countries being common. However, due to the structure of EU, political and fiscal decentralization, the principles of subsidiarity, different legal enforcement and the executive institutions in each member state, the degree to which the EU laws and policies are implemented differs across the EU countries. Farole et al. (2011) note that the individual EU countries often face policy implementation issues related to their own institutional capabilities and governance.

[3] Data for 2022 and 2023 not published by Eurostat at the time of this research.

World Bank records governance indicators for over 200 countries in the world for the period from 1996 to 2022 (as per the current data). The indicators used by the World Bank are Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. Those indicators, in particular the four indictors being Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability, are often used as a proxy for good governance, measuring country's governance quality. The countries such as Finland, Sweden and Denmark mentioned previously have highest cloud computing use in enterprises in the EU, are also some of the top ranked counties globally when it comes to good governance as per the World Bank data and Quality of Government Index (Charron, Lapuente & Dykstra, 2014; Independent, UK paper, issue 2016). Similarly, the countries with lowest cloud computing use in enterprises in the EU, are some of the bottom ranked counties globally when it comes to good governance. Hence, leading to the proposition that good governance impacts cloud computing adoption in the EU enterprises.

In this paper, it is argued that cloud computing adoption in enterprises in individual EU countries is impacted by the specific country's institutional capabilities, those reflecting country's quality of governance. It is posed that the countries with higher quality of governance will have lower corruption level, higher applicability of rule of law and higher voice and accountability, hence in general are more effectively governed. Therefore, businesses operating in those countries will also be able to operate more effectively, with stronger business confidence and lower risk perceptions, focusing on business objectives such as cloud computing adoption rather than being influenced by the presence of corruption and/or lack of the applicability of laws. As such, those countries will have higher cloud computing adoption in enterprises as opposed to those with weaker quality of governance. No relevant research has been found that establishes an empirical link between a country's governance quality and its effects on cloud computing adoption by enterprises, a gap this paper is attempting to address.

The analysis is performed using the regression techniques on the panel data (per EU country/per year) of the percentage use of cloud computing service in enterprises and the average of the World Bank governance indicators, as a measure of good governance in EU countries.

This study aims to provide insights to whether and how institutions reflected in country's governance quality indicators matter when it comes to the cloud computing adoption in the EU enterprises. The findings of the study are important to both the EU government for the purpose of achieving its digitalization targets set in the Industrial Strategy 2030, as well as to the EU firms which gain direct benefits from increased cloud computing adoption in their business.

There are many different perspectives that can be taken in order to study technology adoption such as cloud computing in the EU enterprises, such as socio-economic, demographic, institutional or others, as was done in previous research. This paper, inspired by neo institutionalism as one of the main theories in the organisation research, focuses on the impact of the selected formal institutions reflected in the country's quality of governance on the adoption of cloud computing.

In this article the following structure is applied: Section 2 introduces relevant theory and reflects on the relevant research, followed by the formation of the hypotheses. The section after details the empirical analysis including the description of the data used, method applied and the results obtained. The paper concludes with the discussion, contributions, limitations and implications for research and practice.

2. Literature review and hypotheses development

There has been a number of studies that investigated the digitalization across countries, especially in the EU, such as studies by Tutak and Brodny (2022). In one of their studies, Tutak and Brodny analyse the level of business digital maturity in Europe by investigating the implementation of the digital technologies by companies. The digital technologies their study focused on are artificial intelligence, big data, cloud computing, 3D printing, robotization, the integration of internal process, integration with customers/suppliers, and cybersecurity. They found that Scandinavian countries (Finland and Denmark) and Malta have the highest levels of the implementation of those technologies, while Greece, Romania, Bulgaria, Hungary, and Latvia have the lowest levels. In a different study, Tutak and Brodny (2022)

investigate digital maturity of the SMEs in the EU countries and find that there is a large discrepancy between the initial EU member states (the initial 14) and the new member states (additional 13). They also find that there is a positive relationship between the digital maturity and some economic parameters such as GDP per capita. Both studies by Tutak and Brodny note there is a difference in the technology adoption by the businesses in different EU countries; however, they do not explain the reason for those differences. A study by Kovac, Zmija, Roy, Kusa and Duda (2024) provides a literature review on the topic of digitalisation in the EU and Western Balkan countries and notes that socio-economic context has a potential impact on digitalization, and notes that it has not been sufficiently reflected in the research to date.

Institutions are humanly devised constraints and as such can be either formal or informal (North, 1991). Informal institutions refer to the conventions and norms acceptable (or not acceptable) in the society and are typically reflected in country's national culture. Formal institutions are 'formal rules (constitutions, laws, rights, policies)' (North, 1991, p. 97) and as such are reflections of laws and policies in a country. North (1991) notes that the main purpose of formal institutions is to reduce uncertainty of operating in a certain country. 'Institutions set the boundary conditions under which business occurs (North, 1990), and are the dynamic, sense-making frames that guide actors/organizations to behave in appropriate ways' (Vu, Hartley & Kankanhhalli, 2020).

New institutional theory, or neo institutionalism, study impacts of institutions, both formal and informal, on organisations. According to the new institutional theory (Currie, 2011), 'developing and using technology by organisations is a subject to various institutional constraints or simply institutions.' Karamujic (2025), notes 'the institutions either push or inhibit individuals and groups to adopt the technology and may cause them to behave in not necessarily economic-rationalistic way'. Wan (2005) provides an institution-based theory of firm's strategy and performance, and states that due to the difference in quality of the formal institutions in different countries, resources and capabilities that companies can develop and have access to differ so does their performance.

Governance of a country denotes how the formal institutions are managed on a country level. Country's governance is typically measured by the World Bank governance indicators; hence, for most of countries it is possible to measure the quality of its governance using the World Bank governance indicators. In support of the institutional theory, in this paper we specifically investigate if and how formal institutions measuring the quality of governance (often referred to as good governance) using the World Bank governance indicators impact the adoption of cloud computing by businesses in the EU countries.

There has been increasing focus on good governance and its impact on growth and development of economies and societies. Wu, C-H. (2021) as well as other researchers (Bota-Avram, 2014; Ngobo & Fouda, 2012; Bruno & Claessen, 2010; and others) note that 'country governance is important'. Ngobo and Fouda (2012) based on their study of a relationship between good governance and the performance (profitability) of the firms in 21 African countries, find a positive association between good governance and firms' performance. They note that 'companies located in countries with better governance tend to perform better.' Additionally, they find that 'when the country's income level is lower, an improvement in public governance is more likely to impact a firm's performance than when the income level is relatively higher'.

Bota-Avram (2014) note 'the relevance of governance quality for the ease of doing business, highlighting the necessity to pay enough attention to ensuring good governance in order to provide effective development outcomes.'

Keser and Gokmen (2018) establish a positive relation between governance in the EU and Human Development Index measured a composite of education, literacy and income in that country. They find that at least three governance indicators measured by the World Bank, being Government Effectiveness, Regulatory Quality and Rule of Law have significant positive impact on the development of human capital in the country, hence note that 'better governance in a country, better its human development index.'

Kaufmann and Lafarre (2021) find that country's governance quality is positively associated with sustainability performance, referred to as corporate social performance[4] of the firms in the countries.

Charron, Lapuente and Dykstra (2014) in their research note that 'even in a highly developed area such as Europe, significant quality of governance variation exists between and within countries and that quality of governance has a significant impact on almost all dimensions within societies and economies, such as welfare, education, health, economic performance and others.'

Karamujic (2025), finds that national institutions affect cloud computing adoption; informal institutions such as national culture and few formal institutions such as trade union strength and government effectiveness. The focus was on few formal institutions that were found relevant based on the previous research, with the analysis conducted using the data on 38[5] countries globally.

The study by Baccianti, Labhard and Lehtimaki (2022) investigates the impacts of the quality of institutions on the technology diffusion measured by the number of individual internet users and the number of individual fixed broadband subscriptions in different countries, finding that 'a higher quality of institutions is associated with greater diffusion of digital technologies.'

Using a variety of sources on both national and sub-national level such as World Bank data complemented with variety of survey data, Charron, Lapuente and Rothstein (2011) of the University of Gothenburg developed a novel EQI (European Quality of Government Index), which is now commonly used in the research of country's governance quality impacts. Such one study using EQI is the study by Rodrigues-Pose and Di Cataldo (2014). In their study, Rodrigues-Pose and Di Cataldo assess the impact that the government institutions have on the innovative capacity in the EU, stating that institutions 'define degree of economic uncertainly in the society and the way the collective decisions are made.' Rodrigues-Pose and Di Cataldo note that the individual EU countries' governments vary considerably and argue that the countries with weak government governance will result in lower innovative performance when compared to the countries with stronger government governance settings. Using the European Quality of Government Index (Charron, Lapuente & Rothstein, 2011), and including the World Bank Governance Indicators of Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability, they find that 'the quality of government governance plays a strong and significant role in the innovative potential of European regions.' Rodrigues-Pose and Di Cataldo (2014) additionally find that 'ineffective and corrupt governments represent a fundamental barrier to the innovation in the EU'. Governments that have effective policies and keep corruption under control foster innovation more when compared to their counterparts that are more corrupt and ineffective. Rodrigues-Pose and Di Cataldo (2014) in general acknowledge the importance of institutions on technology adoption and application, noting limited to date academic research of this phenomenon. There are studies that investigate the impacts digitalisation has on good governance by countries and corporates such as its impacts on corruption reduction and the rule of law implementation. However, few studies are found that investigate it the other way around, referring to the impacts the good (country) governance has on digitalisation and technology adoption. Example of such a study is a paper by Baccianti, Labhard and Lehtimaki (2022), centred around the impacts of good (country) governance on the adoption of digital technologies such as internet use and fixed broadband subscription. The research by Baccianti, Labhard and Lehtimaki (2022); however, analyses the impacts on the technology adoption by individuals, not the impact on the technology adoption by organisations and businesses, which is the focus of this study.

Therefore, considering the importance of cloud computing technology adoption to digitalisation, and the lack of research in the countries' quality of governance impacts on cloud computing adoption by the enterprises, this research arose with the specific focus on the EU countries.

The World Bank governance indicators are frequently used as a measure of country's governance quality and are generally accepted to be accurate and reliable (Dollar & Kraay, 2003; Charron, Lapuente

[4] 'Corporate Social Responsibility is the combination of environmental sustainability and social sustainability. Environmental sustainability ensures the long-term stability and resilience of the ecosystems that support human life, whereas social sustainability facilitates the respect and promotion of human rights and other basic social rights'

[5] Some of the EU countries used, such as Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Poland, Portugal, Spain, Sweden

& Dykstra, 2014; Rodrigues-Pose & Di Cataldo, 2014; Rodriguez-Andres, Amavilah & Asongu, 2016; and others). Charron, Lapuente and Rothstein (2011) when working on the European Quality of Government Index identified that four out of six World Bank Governance indicators are particularly relevant when measuring Quality of Government (QoG), being Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability. Those four indicators are also used in this study.

Control of Corruption as per the World Bank definition refers to 'the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.' In corrupt governments, the decisions are not necessarily made for the best of societies and businesses, but rather for individual or privileged groups gains. A number or previous studies identified significant and negative link between corruption and economic growth and development, including technological progress (Rodrigues-Pose & Di Cataldo, 2014). In corrupt organisations and environments, business-oriented and benefiting technology such as cloud computing may not necessarily be supported unless there is a gain for corrupt individuals and groups, which in turn may inhibit the adoption of cloud computing.

Rule of Law as per the World Bank definition refers to 'the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.' Applicability, effectiveness and impartiality of courts are essential for investment and innovation (Rodrigues-Pose & Di Cataldo, 2014; Rose-Acerman, 2001). Vu, Hartley and Kankanhhalli (2020) in their study on country-level predictors of cloud computing adoption find that 'legal systems quality of a country, reflected in a rule of law' positively impact cloud computing.

Government effectiveness index definition refers to 'the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies'. As noted by Rodrigues-Pose and Di Cataldo (2014), effective governments are more capable of developing effective strategies and policies targeting the right areas of investments. Rodriguez-Andres, Amavilah and Asongu (2016) also find that government effectiveness has significant effect on the ICT adoption in their research in Sub Saharan Africa. In my previous research (Karamujic, 2025), in an empirical examination of the cloud computing revenues, as a proxy for cloud computing adoption, over a sample consisting of 38 countries globally (15 EU countries and 23 non-EU) over a five-year period, it was found government effectiveness, as one of six selected national institutions, significantly impacts cloud computing adoption in a country.

Voice and accountability as per the World Bank definition refers to 'the extent to which a country's citizens are able to participate in selecting their government, as well as a freedom of expression, a freedom of association, and a free media.' Rodriguez-Andres, Amavilah and Asongu (2016) relate voice and accountability to political governance in the country, 'the process by which those in authority are selected and replaced'.

As noted earlier, Charron, Lapuente and Rothstein (2011) use Word Bank's governance indicators of Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability as the main pillars and apply combined averages of those four pillars per country to obtain quality of government index. Since those World Bank indicators are only available on a national level, Charron, Lapuente and Rothstein (2011) complement those with survey data on the levels of quality, impartiality and corruption of sub-national governments, resulting in adjusted index on both national and sub-national levels (e.g., Austria as a nation versus sub-national regions of Vienna, Lower Austria, Upper Austria and others). Hence, Charron, Lapuente and Rothstein (2011) obtained adjusted scores which they refer to as European Quality of Government Index (EQI).

As noted before, Rodrigues-Pose and Di Cataldo (2014) in their study on the impact of the government institutions on the innovative capacity in the EU, use the European Quality of Government Index (EQI) as well as separate the World Band good governance indicators of Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability. The methodology applied in Rodrigues-Pose and Di Cataldo (2014) research, even though related to the innovation capacity, can be transposed to other domains, such as digitalization and cloud computing.

Similar to Rodrigues-Pose and Di Cataldo, this study poses that the country's governance quality impacts the use of cloud computing service (instead of the innovative capacity) in the EU. The quality of governance, referring to the quality of the institutions in a country, can affect the speed and the process of adoption of digital technologies. According to Baccianti, Labhard and Lehtimaki (2022), 'the process of digitalisation can benefit from the adjusting of institutions and governance to better suit the introduction of new technologies'. Controlling corruption will result in a fairer assessment of the adoption of technology such as cloud computing on its own merits, rather than being affecting by corrupt behaviours of those involved in the decision process. Hence, higher control of corruption in a country should result in an increased adoption of cloud computing technology. 'An institutional environment with a strong rule of law is likely to increase business confidence and lower risk perceptions of adopting new technology such as cloud computing, as it is known that business interests will be protected in such an environment' (Vu, Hartley & Kankanhhalli, 2020). Hence, higher rule of law applied in a country should result in an increased adoption of cloud computing technology. As noted in my previous paper (Karamujic, 2025), 'in countries with more effective governments, labour markets are more regulated, hence less resistant to implementing labour reducing technologies such as cloud computing. Governments in those countries take an active role in redeploying the workforce to other jobs and reskilling the workforce.' It was also noted that government effectiveness can imply greater application of rules and regulations which could be limiting to cloud computing adoption. Example being GDPR (General Data Protection Regulation) which is in force since 2018 across the EU states. Countries that do not have such regulation may experience greater level of flexibility when it comes to the adoption of cloud computing technology. However, in the context of the EU countries, the regulations such as GDPR are rather strictly imposed and evenly applied across the union, hence not making a significant difference across the member states. Hence, one would expect that higher government effectiveness in an EU country should result in an increased adoption of cloud computing technology. Higher voice and accountability implies higher confidence in political governance in a country, which in general is positive for a business environment, hence one would expect that it will have a positive effect on cloud computing technology adoption too.

Hence overall, one can say that the quality of governance in a country impacts confidence and risk perceptions of adopting a technology such as cloud computing by businesses/enterprises in that country. Thus:

> Hypothesis: The higher the quality of country's governance in a country, the higher cloud computing adoption in the enterprises in that country (focus in this study being the EU).

Where, the quality of country's governance in a country is reflected in the World Band good governance indicators of Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability.

To test the above hypothesis, the dynamic panel regression methods are applied. The methodology and the analysis are further described in Data and Method section.

3. Data and method

3.1. Data

3.1.1. Use of cloud computing services in enterprises in Europe

Eurostat records the data on use of cloud computing service in enterprises in Europe in the period from 2014 until 2021 (with 2019 data missing). The data is based on the results of surveying 148,000 of EU enterprises with at least 10 employees in the EU NACE[6] classification (out of 1.5 million enterprises in this classification in the EU in 2021) including Small-Enterprises (10-49 employees comprising 83%

[6] NACE sectors covers are manufacturing, electricity, gas and steam, water supply, construction, wholesale and retail trades, repair of motor vehicles and motorcycles, transportation and storage, accommodation and food service activities, information and communication, real estate, professional, scientific and technical activities, administrative and support activities, and the repair of computers and communication equipment.

of the data sample), Medium Enterprises (50-249 employees comprising 14% of the data sample) and Large Enterprises (250 or more employees comprising 3% of the sample). The data is expressed as a percentage of enterprises in the NACE sectors (defined earlier in the footnote 4) using cloud computing services. It covers all cloud services being Infrastructure as a Service, Platform as a Service and Software as a Service for all enterprises with 10 employees or more (Source: Eurostat). The data is available for all 27 EU countries in a panel format, per country, per year.

Note: In order to onbtain fully balanced panel, the missing data points were estimated by averaging of the two neighbouring values. For example, if 2015 data for a specific country was missing, it was estimated to be an average of 2014 and 2016 data points. This was only applied in few cases, hence should not have a significant impact on the data. Additionally, the values for 2019 are missing and have been estimated as an average of 2018 and 2020 values.

This data was also used in another research paper that by Karamujic (2025), 'Does Gaia-X foster cloud computing adoption by European (EU) businesses?'

3.1.2. World Bank Governance Indicators (WGI)

Six worldwide governance indicators are Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption.

Each indicator is normalized to a range from -2.5 to 2.5, with zero mean and standard deviation of 1. Higher the value, indicates better perception of a given indicator. The data is captured via various data sources reflecting the perception of governments and is available since 1996 for over 200 countries.

As noted earlier, those indicators are often used as a proxy for the country's governance quality and often referred to as good governance, in particular Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability. The values of those four indicators are used to come to an average per country (per year) as an estimation (proxy) of country's governance quality (good governance) for each of the 27 EU countries. The data range that each indicator takes is somewhat narrower than in the total global data sample, ranging from around -0.5 to 2.5, hence the average of the four indicators ranges from around -0,1 to just below 2.0. Since this research uses logs, 1.0 is added to the average of the World Bank indicators in order to ensure all the values are positive.

3.1.3. European Quality of Government Index (EQI)

The World Bank indicators are measured on a national level and assume that inter-country institutional variations are insignificant. This assumption has been challenged, resulting in a creation of a homogenous European Quality of Government Index (EQI) developed by the University of Gothenburg (Charron, Lapuent & Rothstein, 2011) measured on a sub-national level i.e., per region or province in a country. The index is developed first using an average of four of the World Bank indicators of Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability. Subsequently, those World Bank indices' averages are combined with the survey data on the levels of quality, impartiality and corruption of sub-national governments, resulting in adjusted index on both national and sub-national levels. The values range from around -2.0 to +2.0 with the EU average being 0. Positive values imply regions are above the EU average, while negative values imply regions are below the EU average. Higher the value the better country's governance perception.

The EQI data was first collected in 2010 and then repeated for the periods 2013, 2017 and 2021.

3.1.4. General notes on the data used

All data values are rounded to the third decimal place.

Descriptive statistics and correlations for the data used can found the Appendix, Table A.

3.2. Method

3.2.1. Dynamic regression (GMM)

In this research, the dependent variable is the use of cloud computing services in enterprises in Europe in the period from 2014 until 2021. The independent variable is the country's governance quality (good governance) which is proxied by the average of the four World Bank governance indicators.

It is important to note that the European Quality of Government Index was also considered as a potential proxy for the country's governance; however, it was decided not to be used. The main reason being that the measurement of the European Quality of Government Index is only available for 2 years (2017 and 2021) out of the relevant period from 2014 to 2021 for which cloud computing service in enterprises in Europe data is available. Additionally, the cloud computing service in enterprises in Europe data is only available on national level, with no subnational details, for which the European Quality of Governance Index would be relevant. Moreover, the average of the four World Bank indicators and the European Quality of government Index are highly correlated (0.966, reference Appendix A), hence it is sufficient to use one of them in the analysis, in this case the average of the four World Bank indicators for which yearly data is available for the full period 2014 to 2021.

High correlations present between dependent and independent variable (0.32, reference Appendix, Table A) point to choosing a regression analysis as the main method in this analysis.

Since both dependent and independent variables are available in a panel structure with the observations recorded over the years, panel regression analysis is further selected. Additionally, it is expected that past values of dependent variable, being the use of cloud computing services in enterprises in Europe, affect the future values of the same variable, hence a dynamic regression is used. Vu, Hartley and Kankanhhalli (2020) choose GMM dynamic regression in their study on the analysis of the factors affecting cloud computing. They note that GMM method accounts for the presence of country-fixed effects and is more robust to heteroscedasticity and serial correlation potentially present in the error term compared to other regression techniques. GMM was also used in the study 'The impact of national institutions on cloud computing adoption. Comparison of cloud computing to mobile broadband adoption' (Karamujic, 2025). In that study, the focus was on the impact of informal and some formal institutions on cloud computing adoption in 38 countries globally, using cloud computing revenues as a proxy for the adoption. In this study, the focus is on the impact of the quality of governance in 27 EU countries on the use of cloud computing services in the EU businesses. Additionally, the previous study also compared cloud computing adoption to mobile broadband adoption proxied by the mobile broadband revenues in 48 countries globally. Even though the data and focus of the two studies are different, the methodologies applied are similar.

Log-log regression is also used since useful in interpreting results (one percent change in one of the independent variables (everything else held constant), explains x percent of the variation in the dependent variables). Since the independent variable, the average of the World Bank governance indicators, can take negative values, the data is slightly manipulated in order to apply logs, meaning a fixed value is added to each independent variable to make it positive. See further under Data.

The following statistical model is used to examine the relationship between the use of cloud computing in the EU business and the country's governance quality in the EU countries:

$$\text{lncloud\_use\_enterprise}_{it} = \alpha + \beta\, \text{lncloud\_use\_enterprise}_{it-1} + \gamma X'_{it} + \varepsilon_{it} \qquad \text{Eq (1)}$$

where $\text{lncloud\_use\_enterprise}_{it}$ is the log value of the use of cloud computing services in enterprises in Europe by country i in year t; $\text{lncloud\_use\_enterprise}_{it-1}$ is the lagged level of the log value of the use of cloud computing in the EU enterprises by country i in year t; X'_{it} is a vector containing the log of independent variable, the average of the four World Bank governance indices and a set of control variables as described in Version A, B, C and D below, and ε_{it} is the random error term.

Model versions with varying control variables are:

- *Version A* with net income per capita as a control variable. Net income per capital captures the level of development per country (Source: Worldbank) with the data available for all 27 countries and for each year from 2014 to 2021 time period. As noted in the previous study by Karamujic (2025), 'GMM assumes that the first differences of endogenous variables are uncorrelated with country fixed effects', which could be violated in this data set. Furthermore, the previous research by Ngobo and Fouda (2012), indicated that 'a role of good governance depends upon the country's income level'. To control for this potential violation and bias, net income per capita is added to the analysis.
- *Version B* adds ICT Sector as an additional control variable besides net income per capita. ICT Sector variable captures the percentage of the ICT sector in gross value added, meaning the contribution of the ICT sector to the growth economy (Source: Eurostat), with the data available for all but three EU countries (Cyprus, Ireland and Luxembourg) and the data points available for most of the years for the remaining twenty-four countries. In case the data was missing, the value from a previous year(s) is used (max of three years in the past).
- ICT Sector variable controls for the potential bias that the use of cloud computing is impacted by the extent of ICT sector development in a country.
- *Version C* adds and additional control variable being Broadband internet speed of over 100 Megabits availability in a country (Source: Eurostat) with the data available for all 27 countries and for each year from 2014 to 2021 time period.
- For businesses to effectively use cloud computing services they require internet speed of over 100 Megabits and more (Source: https://www.business.com/internet/bandwidth/), hence this additional variable controls for a potential bias related to the internet speed availability in different countries.
- *Version D* instead of net income per capita, uses GDP per capita in order to capture the level of development per country (Source: Worldbank) which in some studies (Tutak and Brodny, 2022) was found to be significantly related to the digitalization maturity. The rest of the control variables are consistent with the model version C, being ICT Sector and Broadband internet speed.

All model versions use time dummies to control for time-related shocks.

4. RESULTS

Tables 1 and 2 below detail the estimation results for the models specified in Eq (1) resulting from dynamic regression using one step GMM.

TABLE1.
The impact of the average of the four World Bank governance indicators on the use of cloud computing services in enterprises in Europe as per the dynamic regression (GMM one step) from STATA

The impact of the average of the four World Bank governance indicators on the use of cloud computing services in EU enterprises				
	Model A	**Model B**	**Model C**	**Model D**
Use of cloud computing in enterprises _lag value P value [95% Conf. Interval] (Std. Error)	0.829*** 0.000 [0.698 0.959] (0.067)	0.873*** 0.000 [0.706 1.040] (0.085)	0.886*** 0.000 [0.709 1.063] (0.090)	0.892*** 0.000 [0.712 1.072] (0.092)
Average of the four World Bank governance indicators P value [95% Conf. Interval] (Std. Error)	0.274*** 0.000 [0.125 0.424] (0.076)	0.205*** 0.001 [0.079 0.330] (0.064)	0.167** 0.043 [0.006 0.328] (0.082)	0.178** 0.024 [0.023 0.333] (0.079)
Net Income Per Capita	-0.000000678	-0.000000511	-0.00000053	
P value [95% Conf. Interval] (Std. Error)	0.164 [-0.00000163 0.000000278] (0.000000488)	0.575 [-0.0000023 0.00000128] (0.000000913)	0.563 [-0.00000233 0.00000127] (0.000000916)	
GDP Per Capita P value [95% Conf. Interval] (Std. Error)				-0.000000634 0.425 [-0.00000219 0.000000923] (0.000000795)
ICT Sector P value [95% Conf. Interval] (Std. Error)		0.004 0.368 [-0.005 0.013] (0.005)	0.002 0.786 [-0.010 0.013] (0.006)	0.001 0.882 [-0.010 0.013] (0.006)
Broadband P value [95% Conf. Interval] (Std. Error)			0.0003 0.213 [-0.0002 0.0009] (0.0003)	0.0003 0.209 [-0.0002 0.0009] (0.0003)
Constant P value [95% Conf. Interval] (Std. Error)	0.254*** 0.001 [0.105 0.402] (0.076)	-36.717*** 0.006 [-62.804 -10.631] (13.310)	-32.147** 0.038 [-62.536 -1.756] (15.505)	-31.95** 0.035 [-61.652 -2.251] (15.153)
N	189	166	166	166
Number of groups	27	24	24	24
Observations per group	7	5	5	5

Numbers rounded to 3 decimal places; *p < 0.1, **p < 0.05; *** p < 0.01.

TABLE 2.
GMM Hypotheses tests

The impact of European Quality of government Index on the use of cloud computing services in EU enterprises				
	Model A	**Model B**	**Model C**	**Model D**
Arellano-Bond test for AR(2)	z = 0.9 Pr > z = 0.367	z = 1.04 Pr > z = 0.297	z = 1.02 Pr > z = 0.308	z = 1.02 Pr > z = 0.307
Hansen test of overid. restriction	chi2(26) = 17.72 Prob > chi2= 0.886	chi2(26) = 16.46 Prob > chi2= 0.924	chi2(26) = 13.41 Prob > chi2= 0.980	chi2(26) = 11.68 Prob > chi2= 0.993

Firstly, important to note that Arellano-Bond tests and Hansen tests in Table 2 above, confirm the validity of GMM used, hence we can proceed with the analysis of the results in Table 1.

In Table 1, as expected, the lag values of the dependent variable are significant in all four models A, B, C and D with p values of 0.000, hence the past values of the use of cloud computing in the EU enterprises do significantly affect the future values.

This research's main hypothesis poses that the higher the governance quality in an EU country (proxied by the average of the four World Bank governance indicators), the higher cloud computing adoption in the enterprises in that EU country. As can be seen in all four models, this hypothesis is supported (Model A with p value of 0.000, Model B with p value of 0.001, Model C with p value of 0.043 and Model D with p value of 0.024). The relationship between the governance quality in EU countries and the use of cloud computing by EU enterprises is positive and significant indicating that 1% increase in the governance quality in the EU countries as measured by the average of the four World Bank governance indicators results in about 0.2% increase in cloud computing adoption in the enterprises in that EU country (measured as a percentage of EU enterprises that adopt cloud computing in that country). More accurately, 1% increase in the governance quality in the EU countries results in 0.274% increase in the percentage of EU enterprises that adopt cloud computing in that country as per Model A, 0.205% increase as per Model B, 0.167% as per Model C and 0.178% as per model D. Hence, the results obtained seems to be consistent across different models used with different control variables applied.

Model A has the most data points, while models B, C and D have less data since ICT Sector information is missing for 3 countries in the data sample.

5. DISCUSSION, CONCLUSION, CONTRIBUTION, AND LIMITATIONS

In this research, it was illustrated that the country's governance quality (good governance) positively affects the use of cloud computing in the EU enterprises. The empirical analysis was conducted using the average of the four World Bank governance indicators (Control of Corruption, Rule of Law, Government Effectiveness and Voice and Accountability) as a proxy for the governance quality. To accommodate for potential biases, four different models were used to control for factors such as net income, GDP, ICT sector development as well as broadband internet availability in different countries. All four models tested show the consistent results in support of the research hypothesis: The higher the governance quality in an EU country, the higher cloud computing adoption in the enterprises in that EU country.

5.1. CONTRIBUTIONS

This research makes several important contributions. Firstly, it supports the exiting theories, in particular neo institutional theory. According to the neo institutional theory, developing and using technology by organisations is a subject to various formal institutions. In this study we show that formal institutions represented by the World Bank governance indicators, being Control of Corruption, Rule of

Law, Government Effectiveness and Voice and Accountability are influencing the use of cloud computing services in the EU enterprises. As such, this study contributes to the body of research that supports the neo institutional theory, as one of the main theories in the organisation research. As noted in my previous research, 'In support of the neo-institutional theory, national institutions may cause individuals in a certain national setting not to behave in a necessarily economic-rationalistic way; hence, the success of the technology implementation, such as cloud computing, will depend on the organizations understanding of that national setting and taking right measures to manage it in the best possible way' (Karamujic, 2025).

Secondly, it empirically shows how the country's governance quality impacts the use of cloud computing services in the EU enterprises, stating that higher the governance quality in an EU country, the higher cloud computing adoption in the enterprises in that EU country. Similar research has not been conducted year to date and provides an important finding for both researchers and practitioners. Researchers can further investigate this topic and possibly extend it to different geographies and additional institutions, while practitioners, both governments and businesses, can take it as an input it its planning and management, as described further under practical implications.

Thirdly, this research contributes to the growing literature on the importance of good governance and its impact on growth and development of economies and societies.

5.2. Practical implications

Adoption of cloud computing by enterprises is one of the main objectives of the EU Commission Industrial Strategy 2030 with the aim of accelerating the digitalisation and transforming the EU into 'the continent with a high share of digitalised business.' The goal of 75% of the EU enterprises to have taken cloud computing services[7] by 2030 is quite a challenge, considering the 2021 EU average was just over 40%, with very large variations across the countries. This research aims to provide additional input to the EU government(s) and decision makers and support them with additional insights in terms of drawing their attention to the importance of the country's governance quality, and the improvement of it, for continued growth in the use of cloud computing in the EU enterprises, hence increase in digitalisation and achievement of their strategic objectives. Improving control of corruption, hence reducing corruption in a country; implementing greater application of the rule of law, hence ensuring legal system works well; ensuring government and its institutions work effectively and independently are all very important from various aspects, and in this research, it is shown that they are also very important for a country's adoption of technology such as could computing. EU and individual country governments and policy makers have significant influence on all of those as they are formal institutions and reflect policies and laws in a country. 'When aiming to increase technology adoption on a country level, the government should look at its laws and policies more critically and if necessary, adjust in order to enable higher adoption success' (Karamujic, 2025).

This research also aims to provide additional input to the EU businesses, for which the adoption of cloud computing and acceleration of digitalisation is essential given its many cost and performance business benefits. Enterprises operating in countries with higher governance quality are more likely to implement cloud computing in their operations. Hence, when establishing or planning to grow in certain EU countries, this research provides important insights to EU businesses.

This research also confirms previous research (Wan, 2005; Ngobo & Fouda, 2012) that companies located in countries with better governance tend to perform better, in this case measured by their adoption of cost and performance benefiting technology such as cloud computing.

5.3. Limitations

This research is focused on the use of cloud computing services in the EU enterprises and should be used in the context of the EU. In would be interesting to conduction similar research on a wider geography; however, Eurostat only records the use of cloud computing services in the EU enterprises, and no similar

[7] Cloud computing is one of the key technologies enabling digital transformation. The other digital technologies mentioned in the Commission Industrial Strategy 2030 are Big Data and AI (Artificial Intelligence)

data is available on a wider data sample. Additionally, there are also other aspects and parameters that may be considered relevant for cloud computing adoption; however, they are omitted and are not focus of this study. Focus is on the institutions echoed in country's governance quality indicators as previously explained.

References

Baccianti, C., Labhard, V., & Lehtimaki, J. (2022). Digitalisation, institutions and governance, and diffusion: mechanism and evidence. *ECB Working Paper Series No.2022/2675.*

Bota-Avram, C. (2014). Good governance and doing business: Evidence from a cross-country survey. *Transylvanian Review of Administrative Sciences, 41*, 27-45.

Bruno, V., & Claessens, S. (2010). Corporate governance and regulation: Can there be too much of a good thing? *Journal of Financial Intermediation*, *19*, 461-82.

Charron, N., Lapuente V., & Rothstein, B. (2011). *Measuring quality of government and sub national variation*. Report for the EU Commission of Regional Development, European Commission, Directorate-General Regional Policy.

Charron, N., Lapuente, V., & Dykstra, L. (2014). Regional governance matters: Quality of Government within European Union Member States. *Regional Studies*, *48*, 68-90.

Charron, N., Lapuente, V., Bauhr, M., & Annoni, P. (2022). Change and Continuity in Quality of Government: Trends in subnational quality of government in EU member states. *Investigaciones Regionales - Journal of Regional Research*, *53*, 5-23.

Charron, N., Dijkstra, L., & Lapuente, V. (2015). Mapping the regional divide in Europe: A measure for assessing quality of government in 206 European regions: *Social indicator research, 122*(2), 315-346.

Charron, N., Lapuente, V., & Annoni, P. (2019). Measuring Quality of Government in EU Regions Across Space and Time. *Papers in Regional Science.*

Chen, T., Chuang, T., & Nakatani, K. (2016). The Perceived Benefits Business Benefit of Cloud Computing: An Exploratory Study. *Journal of International Technology and Information Management*, *25*(4), 100-122.

Currie, W.L., & Galliers, R.D. (2011). *The Oxford Handbook of Management Information Systems: Critical Perspectives and New Directions*. Oxford University Press.

Dollar, D., & Kraay. A. (2003). Institutions, trade and growth. *Journal of Monetary Economics*, *50*, 133-162.

European Quality of Government Index (EQI). (2024). Retrieved April 14, 2024, from https://www.gu.se/en/quality-government/qog-data/data-downloads/european-quality-of-government-index

European Commission. (2023). *2030 Digital Compass: the European way for the Digital Decade.* Retrieved September 11, 2023, from https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52021DC0118

Eurostat Database. (2022). *Cloud computing services by size class of enterprise.* Retrieved October 15, 2022, from https://ec.europa.eu/eurostat/databrowser/view/isoc_cicce_use/default/table?lang=en

Eurostat. (2023). *Cloud Computing – statistics on the use by enterprises.* Retrieved June 11, 2023, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Cloud_computing_-_statistics_on_the_use_by_enterprises&stable=0&redirect=no#Enterprises_using_cloud_computing

Eurostat. (2024). *Broadband internet coverage by speed.* Retrieved July 3, 2024, from https://ec.europa.eu/eurostat/databrowser/view/isoc_cbs/default/table?lang=en&category=isoc.isoc_ad.isoc_c

Eurostat. (2024). *Percentage of the ICT sector in Gross value added.* Retrieved July 3, 2024, from https://ec.europa.eu/eurostat/databrowser/view/isoc_bde15ag/default/table?lang=en&category=isoc.isoc_se

Farole, T., Rodrıguez-Pose, A., & Storper, M. (2011). Cohesion policy in the European Union: growth, geography, institutions. *Journal of Common Market Studies*, *49*, 1089-1111.

Independent. *The 17 best-governed countries in the world.* Issue 2016. Retrieved April 7, 2024, from https://www.independent.co.uk/news/world/politics/17-bestgoverned-countries-world-a7417096.html

Kaufmann, W., & Lafarre, A. (2021). Does good governance mean better corporate social performance? A comparative study of OECD countries. *International Public Management Journal, 24*(6), 762-791.

Karamujic, L. (2025). Impact of national institutions on cloud computing adoption. Comparison of cloud computing to mobile broadband adoption. *Journal of Global Information Technology Management, 28*(1), 6-29.

Karamujic, L. (2025). Does Gaia-X foster cloud computing adoption by European (EU) businesses? *Under review by Journal of the Knowledge Economy.*

Keser, A., & Gokmen, Y. (2018). Governance and Human Development: The Impacts of Governance Indicators on Human Development. *Journal of Public Administration and Governance, 8*(1), 26-39.

Kovac, N., Zmija, K., Roy, J. K., Kusa, R., & Duda, J. (2024). Digital divide and digitalization in Europe: A bibliometric analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *19*(2), 463-520.

Ngobo, P.V., & Fouda, M. (2012). Is 'Good' governance good for business? A cross-national analysis of firms in African countries. *Journal of World Business*, *47*, 435-49.

North, D.C. (1990). *Institutions, institutional change and economic performance.* Cambridge University Press.

North, D.C. (1991). Institutions. *Journal of Economic Perspectives*, *5*(1), 97-112.

Plummer, D.C., et al. (2008). Cloud computing: Defining and describing an emerging phenomenon. Gartner, June 17, 2008.

Rodriguez-Andres, A., Amavilah, V., & Asongu, S. (2016*). Linkages between Formal Institutions, ICT Adoption and Inclusive Human Development in Sub Saharan Africa*. 2016 African Governance and Development Institute, AGDU Working Paper, No. WP/16/026.

Rodriguez-Pose, A., & Di Cataldo, M. (2014). Quality of government and innovative performance in the regions in Europe. *Journal of Economic Geography*, *15*, 673-706

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*, 71-102.

Romer, P. M. (1994). The Origins of Endogenous Growth. *The Journal of Economic Perspectives, 8*(1), 3-22.

Rose-Ackerman, S. (2001). Trust, honesty, and corruption: reflection of the state-building process. *European Journal of Sociology*, *42*, 27-71.

Shera, A. (2011). Corruption and the Impact on the Economic Growth. *Journal of Information Technology & Economic Development, 2*(1), 39.

Tutak, M., & Brodny, W. (2022). Business Digital Maturity in Europe and Its Implication for Open Innovation. *Journal of Open Innovation: Technology, Market, and Complexity, 8*(27), 1-23.

Tutak, M., & Brodny, W. (2022). Digitalization of Small and Medium-Sized Enterprises and Economic Growth. Evidence for the EU-27 Countries. *Journal of Open Innovation: Technology, Market, and Complexity, 8*(2), 67.

Vithayathil, J. (2017). Will cloud computing make the Information Technology (IT) department obsolete? *Information Systems Journal*, *28*(4), 634-649.

Vu, K., Hartley, K., & Kankanhalli, A. (2020). Predictors of Cloud Computing Adoption: a Cross-Country Study. *Telematics and Informatics*, *Vol. 52*.

Wan, W. P. (2005). Country resource environments, firm capabilities, and corporate diversification strategies. *Journal of Management Studies*, *42*(1), 161-182.

Worldbank. *Worldwide Governance Indicators.* Retrieved February 25, 2024, from https://www.worldbank.org/en/publication/worldwide-governance-indicators

Worldbank. *Adjusted net income per capita in USD.* Retrieved April 21, 2024, from https://data.worldbank.org/indicator/NY.ADJ.NNTY.PC.CD

Worldbank, *GPD per capita.* Retrieved July 3, 2024, from https://data.worldbank.org/indicator/ny.gdp.pcap.cd?most_recent_value_desc=false

Worldometer, *National GDP data 2014-2011*. Retrieved July 3, 2024, from https://www.worldometers.info/gdp/

Wu, C-H. (2021). Article on the Moderating Effects of Country Governance on the Relationship between Corporate Governance and Firm Performance. *Journal of Risk and Financial Management, 14*(3), 140.

Xue, C.T.S., & Xin, F.T.W. (2016). Benefits and Challenges of the Adoption of Cloud Computing in Business. *International Journal on Cloud Computing: Services and Architectures*, *6*(6).

ORCID

Lejla Karamujic https://orcid.org/0009-0003-3411-9093

Appendices

Table A.
Descriptive statistics and bivariate correlations

Variable Name	N	Mean	Standard Deviation	Min	Max	1.	2.	3.
1.Use of cloud computing services in enterprises in Europe	216	28.758	16.645	4.9	75.5	1	0.32***	0.35***
2.Average of the four World Bank governance indicators	216	1.026	0.555	-0.1	1.963	0.32***	1	0.966***
3.European Quality of government Index (EQI)	216	0.38	0.948	-1.812	1.753	0.35***	0.966***	1

Numbers rounded to 3 decimal places; *p < 0.1, **p < 0.05; *** p < 0.01.

Junta Directiva

Presidente: Fernando Rubiera Morollón
Secretario: Rosina Moreno Serrano
Tesorero: Vicente Budí Orduña

Vocales:

André Carrascal Incera (Comisión Ejecutiva)
Ángeles Gayoso Rico (Comisión Ejecutiva)
Juan de Lucio Fernández (Comisión Ejecutiva)
María José Murgui García (Comisión Ejecutiva)
Juan Carlos Rodríguez Cohard (Comisión Ejecutiva)

José Antonio Camacho Ballesta (A. Andaluza)
Jaime Vallés Jimenéz (A.Aragonesa)
Ana Viñuela Jiménez (A. Asturiana)
Adolfo Maza Fernández (A. Cántabra)
José Manuel Díez Modino (A. Castellano-Leonesa)
Agustín Pablo Álvarez Herranz (A. Castellano-Manchega)
Àlex Costa Sáenz de San Pedro (A. Catalana)
Alberto Franco Solís (A. Extremeña)
Xesús Pereira López (A. Gallega)
Raúl Mínguez Fuentes (A. Madrileña)
José Antonio Illán Monreal (A. Murciana)
Luisa Alamá Sabater (A. Valenciana)
Amaia Altuzarra Artola (A.Vasca y Navarra)

La AECR forma parte de la ERSA (European Regional Science Association) y asimismo de la RSAI (Regional Science Association International).
Sus objetivos fundamentales son:

- Promover la Ciencia Regional como materia teórica y aplicada al territorio proveniente de la confluencia sobre el mismo de disciplinas y campos científicos diferentes que contribuyan a un desarrollo armónico y equilibrado del hombre, medio y territorio.
- Crear un foro de intercambio de experiencias favoreciendo la investigación y difusión de métodos, técnicas e instrumentos que afecten a la Ciencia Regional.
- Promover relaciones e intercambios a nivel internacional sobre Ciencia Regional.
- Impulsar el estudio de la Ciencia Regional en los centros docentes y de investigación.
- Promover publicaciones, conferencias y cualquier otra actividad que reviertan en una mejora del análisis y las acciones regionales.
- Colaborar con la Administración Pública, a todos los niveles, para una mejor consecución de los fines de la asociación y el desarrollo del Estado de las Autonomías.
- La asistencia técnica a la Administración Pública u otras instituciones, públicas o privadas, así como a la cooperación internacional en el ámbito de sus objetivos.

Más información:

Conxita Rodríguez i Izquierdo
Teléfono y Fax: +34 93 310 11 12 - E-mail: info@aecr.org
Página web: www.aecr.org

Este número ha sido patrocinado por la **Dirección General de Fondos Europeos** y cofinanciado por el **FEDER** (Fondo Europeo de Desarrollo Regional)

GOBIERNO DE ESPAÑA MINISTERIO DE HACIENDA

Unión Europea

"Una manera de hacer Europa"

Journal of Regional Research
Investigaciones Regionales

Asociación Española de Ciencia Regional
C/ Viladomat, 321, entresuelo 08029 Barcelona
Teléfono y Fax: +34 93 310 11 12
E-mail: info@aecr.org www.aecr.org