

# Geographical distribution of the COVID-19 pandemic and key determinants: Evolution across waves in Spain

Rosina Moreno  | Esther Vayá

AQR-IREA, University of Barcelona,  
Barcelona, Spain

## Correspondence

Rosina Moreno, AQR-IREA, University of  
Barcelona, Barcelona, Spain.  
Email: [moreno@ub.edu](mailto:moreno@ub.edu)

## Funding information

Agència de Gestió d'Ajuts Universitaris i de  
Recerca (AGAUR), Generalitat de Catalunya,  
Grant/Award Number: 2020PANDE00060

## Abstract

Amidst the COVID-19 pandemic, most research has examined specific temporal snapshots. This study diverges by offering a comprehensive analysis of COVID-19 incidence across the Spanish provinces throughout six distinct waves of the pandemic. Using *spatial* exploratory techniques, we find no single pandemic; rather, there have been waves. Significant differences in the spatial distribution of cases and deaths across six waves show that each has unique characteristics. Homogeneous conclusions cannot be drawn at the national level. Notable regional differences in the pandemic's spatial distribution suggest a need for subnational responses, reflecting variations in climate, economic dynamism, sectoral specialisation, and socio-health resources. Spatial regression models show that the main determinants of COVID-19 incidence depend on stage. Traditional factors commonly associated with epidemiological studies, such as temperature, exerted significant influence during the pandemic's onset. However, as mobility restrictions were enforced and vaccination campaigns were rolled out, economic conditions, and especially levels of economic activity, emerged as increasingly significant determinants.

## KEYWORDS

climate, COVID-19 pandemic, exploratory spatial data analysis, pandemic's diffusion, Spanish provinces, spatial regressions

## 1 | INTRODUCTION

From the end of 2019, when the SARS-CoV2 virus was detected and the COVID-19 outbreak appeared in Wuhan, China, to mid-June 2022,<sup>1</sup> more than five billion people have been infected, and more than six million people have died due to this world-wide pandemic. According to the International Monetary Fund (2020), the average global gross domestic product (GDP) dropped by 3.9% from 2019 to 2020, making it the worst economic downturn since the Great Depression. The pandemic outbreak caught populations and governments off guard, resulting in both the collapse or the near collapse of the health system worldwide and a slew of containment and mitigation measures,

such as the obligation to wear masks in inner spaces, increased testing, contact tracing, lengthy lockdowns, quarantines, and mobility restrictions. Those had an uneven impact on the spread of the virus.

Even within Europe, in the first six months of 2020, Belgium, Italy, Spain, and the United Kingdom had the greatest incidence relative to their populations, whereas many Central and Eastern European nations were spared. Deaths in the Baltic States, Bulgaria, Czechia, Hungary, and Slovakia were lower than the previous five-year average. In the first half of 2020, Denmark and Germany similarly had no total excess mortality (Statista, 2023).

The economic literature analysing the pandemic has been profuse in the last two years, dealing to a

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Geographical Research* published by John Wiley & Sons Australia, Ltd on behalf of Institute of Australian Geographers.

greater extent with the study of its determinants as well as its impact (for a literature review, see Brodeur et al., 2021). More specifically, some authors have shown the relevance of geography to better understand the COVID-19 crisis, and several works have focused on the spatial distribution of the pandemic and the determinants behind the heterogeneity in such patterns (see, for instance, Bissell, 2021; Burton, 2021). In addition, some authors have analysed the temporal evolution of the pandemic by focusing on specific moments, primarily some days or weeks in the first wave of the pandemic (see, for instance, Briz-Redón & Serrano-Aroca, 2020; Wang et al., 2021).

The evolution of COVID-19 has been deeply intertwined with those using the theory of diffusion, who seek to understand how infectious diseases spread through populations and across geographic territories (Chang et al., 2021; Elliott & Wartenberg, 2004). We adhere to the idea that the pandemic follows a dynamic evolution due to the different conditions and circumstances in each wave, so that the conclusions obtained for specific points in time may not be directly transferable to other moments of the pandemic.

We focus on the Spanish case, a decision justified by a dual rationale. First, Spain experienced prolonged and significant impacts from the pandemic, exhibiting pronounced geographical variations within its territory. Second, the governance of the state of alert and decisions regarding the implementation or easing of restrictions were decentralised to regional governments. We gain deep insight from an exploration of the spatial distribution of the pandemic in the Spanish provinces across the six waves, from the beginning of the pandemic until March 2022. We compare the pandemic's evolution across waves, while considering the changes in the significance of the determinants of the pandemic's spread across them. We use spatial exploratory techniques as the method to study the spatial pattern of the incidence of the pandemic and take full account of this spatial dimension in the regressions through the estimation of spatial autoregressive models.

The outline of the paper is as follows. Section 2 provides an overview of the literature. In Section 3, we describe the data, whereas Section 4 uses a spatial exploratory analysis to offer the main patterns of the spatial distribution of the pandemic across the Spanish provinces across the six waves. Section 5 presents the main findings on the determinants of such spatial distribution through regression analyses, and Section 6 states conclusions.

## 2 | LITERATURE REVIEW

### 2.1 | Review of research on the spatial distribution of the pandemic

Given the relevance of the COVID-19 pandemic, there is an extensive body of literature studying the dynamics

#### Key insights

A detailed analysis of COVID-19 reveals multiple pandemics across different waves, each with its own characteristics and determining factors. Highly relevant at the pandemic's onset, when the first state of emergency was declared, were factors such as temperature, which are significant in epidemiological studies. However, after mobility restrictions and widespread vaccination processes were implemented, economic conditions became more significant. Regional variations and the importance of the geographic scale in the analysis highlight the need for tailored and context-specific responses, ensuring more effective and efficient management of the health crisis.

of the spread of the virus and its effects on the economy, even if research started only about two years ago. We provide a brief review of the vast literature on the topic, focused first on those papers that place a special emphasis on the spatial/territorial characteristics of COVID-19 and, second, on those dealing with our case study (that is, the Spanish case).

Even though findings from prior literature concerning the spatial dispersion of the disease are specific to the areas and time periods under consideration, several issues are worth noting. One study by Amdaoud et al. (2021) used spatial models to analyse the heterogeneity of the spread of the COVID-19 pandemic across 125 regions in 12 European countries. The authors found that spatial clusters existed and that income and public health policies explained disparities across regions. In another study, Sun et al. (2021) found a greater spatial inequality in the United Kingdom in COVID-19 mortality than in mortality from other causes. Cao et al. (2020) examined the case-fatality rate in 209 countries and territories worldwide and found a significant correlation with population size, which may imply a strain on healthcare and lower treatment efficiency in countries with large populations. Ehlert (2021) found evidence that the infections and deaths in 401 German counties to June 2020 were positively and significantly related to median age, the number of people working in care of the elderly, early cases since the beginning of the pandemic, and population density. For New York City, Yang et al. (2021) found that COVID-19 case rates were positively related to racial minority groups, household size, and the elderly population, while they were negatively related to the number of teleworkers. Fonseca-Rodríguez et al. (2021) conducted a study in Sweden and found that the virus was associated with population density, the proportion of immigrants, and the proportion of people 65+ years old.<sup>2</sup>

The methods used in these studies include spatial exploratory techniques, basically through disease mapping and the identification of spatial clusters. In addition, the studies tended to estimate spatial models to analyse the determinants of the spatial pattern observed (Elhorst, 2014), all of them implicitly considering the presence of spatial autocorrelation in the estimation (for a review of papers using spatial techniques, see Fatima et al., 2021).

For the Spanish case and with special emphasis on the spatial component of the pandemic, Maza and Hierro (2022) focused on the distribution of COVID-19 among municipalities in Madrid during the first wave of the pandemic, finding that those territories with more mobility as well as those with a higher level of tourism had a higher incidence. Along the same line, Hierro and Maza (2023) suggested an adapted spatial Markov chain methodology that involves estimating both an unconditional and a conditional spatial contagion index. With respect to the primary drivers of spatial contagion, the authors concluded that elevated intermunicipal mobility before confinement served as a catalyst for the development of positive spatial dependence in subsequent cumulative incidence rates. Additionally, the conditional spatial contagion index illustrated that densely populated municipalities with sizable immigrant populations were the most susceptible to spatial contagion dynamics during the early stages of the pandemic in Madrid.

Gullón et al. (2022) found evidence for Madrid that there is an unequal distribution of COVID-19 incidence by neighbourhood deprivation (March 2020–September 2021), whereas González et al. (2022) obtained evidence that there is spatial correlation in the distribution of the incidence of the pandemic across Spanish regions and that both socioeconomic variables as well as those of spatial interaction are significant. Páez et al. (2021) used spatial SURE models to investigate Spanish provinces from March to April 2020 and concluded that a higher incidence of the disease is associated with a higher GDP per capita and the presence of mass transit systems, lower population density, and a higher percentage of older adults. Romero and Arroyo (2022) found that in the period March 2020–January 2021, the pandemic was especially suffered in urban areas with higher population density and higher levels of contamination, and Briz-Redón and Serrano-Aroca (2020) failed to obtain consistent evidence of a relationship between the accumulated number of cases in the provinces of Spain and temperature values between February and March 2020. Orea and Álvarez (2022) investigated the spread of COVID-19 throughout the provinces of Spain and evaluated the efficacy of the nationwide lockdown imposed on 14 March 2020 in combating the pandemic. Researchers accomplish these goals by employing a spatial econometric model, which offers an alternative approach to the widely used, reproduction-based models found in epidemiological literature.

## 2.2 | Main determinants of the diffusion of COVID-19

Because of the heterogeneous geographical spread of COVID-19, researchers from all over the world have been investigating the spatial distribution of the disease in conjunction with its main determinants. In accordance with the findings of the literature succinctly surveyed in Section 2.1, we can classify the main determinants into five dimensions, namely, climatology, demographic factors, agglomeration, connectivity, and economic factors.

### 2.2.1 | Climatology

Viruses transmit more easily depending on climate conditions, such as average temperatures and rainfall (Dalziel et al., 2018). Previous studies have suggested a correlation between weather and the COVID-19 pandemic that is similar to that of other viral infectious diseases, such as influenza (Ficetola & Rubolini, 2021; Ma, Lai, et al., 2020; Ma, Zhao, et al., 2020; Tosepu et al., 2020). According to Wang et al. (2020) and Sajadi et al. (2020), the climatic characteristics of the areas in which the incidence of COVID-19 was higher were average temperatures between 5°C and 11°C and relative humidity levels between 50% and 70%. According to this literature, an increase in temperatures and air humidity levels associated with the arrival of spring in the northern hemisphere could significantly reduce the transmission and spread of the coronavirus. In contradiction, other studies have found that meteorological conditions may not be associated with COVID-19 in terms of absolute humidity (Shi et al., 2020) or temperature (Jamil et al., 2020; Xie & Zhu, 2020). According to the authors, the previous results that showed evidence for a correlation between meteorological factors and COVID-19 transmission were likely to be an artefact, reflecting the pathways of the spread; that is, several of the previous studies in favour of the role of climate were performed considering only meteorological factors, without accounting for nonmeteorological variables that might also be decisive.

### 2.2.2 | Demography

Initially, higher incidences were attributed to ageing. Several studies have reported age and underlying diseases as the most important risk factors for death by COVID-19 (Liu et al., 2020; Morley & Vellas, 2020; Onder et al., 2020). Iyanda et al. (2022) conducted a broad study in the United States to explore the health and social determinants of the spread of COVID-19 and found that age plays a central role in determining the spread over all the space. Because

Spain has for decades been ranked among the top 10 countries in the world for highest life expectancy, this is an issue to consider in the Spanish case. It is possible that the ageing population in Spain could be behind the remarkably hard impact of the disease in this country.

### 2.2.3 | Agglomeration

Population density was accounted for in most of the studies analysing the factors that influenced the spread of COVID-19 with the observation that it is a telling factor for risk exposure (Ahmadi et al., 2020; Bayode et al., 2022; Bhadra et al., 2021; Coccia, 2020; Ehler, 2021; González et al., 2022; Mollalo et al., 2020; Pequeno et al., 2020; Wong & Li, 2020). Large and dense European cities were regarded as the focus of the spread of the coronavirus, and most of the literature has concluded in this direction. Even though density may have shaped the early outbreaks, it did not seem to influence its related mortality over time (Carozzi et al., 2020). For instance, Páez et al. (2021) found a negative influence of population density in Spanish provinces.

### 2.2.4 | Connectivity

Linked to large agglomerations, we may think that it is not only relevant to consider the influence of population density but also its greater connectivity (Coelho et al., 2020). Highly connected cities were among the first to be hit by the virus. This connectivity can take place with extra-regional agents as well as internally. From an extra-regional connectivity perspective, the World Tourism Organization (2019) reported that Spain was the country with the second highest international tourist arrivals in the world in 2019. For this reason, we consider that the entry of tourists can be extremely significant in trying to estimate the causes of COVID-19 diffusion in Spain (in line with Maza & Hierro, 2022). As for connectivity within a region or even in cities, public transportation has been related to the spread of contagious diseases (Wang et al., 2020). Páez et al. (2021) reported that the presence of mass transit systems in a province implies a clear positive impact on the diffusion of the disease.

### 2.2.5 | Wealth

A further determinant of the diffusion of the virus is economic wealth, although there is no consensus in the sign of its effect. Wealthier regions may present a lower diffusion of the illness because they tend to concentrate more activities that produce nontraded goods, which

would imply that wealthier individuals remain more active even during lockdowns. In contrast, infectious diseases would tend to have a greater effect on the poorest neighbourhoods because they face more challenges in maintaining social distancing than wealthier individuals and have less access to resources to reduce the chances of infection. In addition, poorer individuals may lack access to medical services and basic resources for living. Less wealthy areas also have a higher proportion of workers in manual occupations who cannot telework and have more difficulty in complying with shelter-in-place orders (Almagro & Orane-Hutchinson, 2020). Accordingly, Baena-Díez et al. (2020) obtained evidence of a negative relationship between income level and COVID-19 incidence for the different districts in the city of Barcelona. Mena et al. (2021) similarly found that mortality rates of young people in Chile were lower in high-income municipalities than in low-income municipalities. However, greater economic capacity could promote greater physical mobility, a factor that spreads the virus. As highlighted by Gong and Zhao (2022), COVID-19 in some European countries was imported by comparatively wealthy travellers, such as Chinese entrepreneurs and ski tourists from the Alps. In this respect, Amdaoud et al. (2021) obtained a positive and significant relationship in European regions between the level of GDP per capita and the level of COVID-19 death rates.

Although not considered in our paper,<sup>3</sup> other determinants that could explain the uneven geography of the COVID-19 pandemic are institutional factors, for instance, formal institutional quality across European regions may imply different capacities to effectively implement measures to prevent and combat the pandemic; sociological factors, such as the tendency to meet with friends and family in celebrations (Rodríguez-Pose & Burlina, 2021) and the vaccination level in the last three waves; and differences in health systems that can influence the capacity to detect and treat outbreaks (Ahmed et al., 2020; Bauer et al., 2020; Liang et al., 2020).

## 3 | DATA AND DESCRIPTIVES: THE GEOGRAPHY OF SIX WAVES

### 3.1 | Data and variables

#### 3.1.1 | Incidence of the COVID-19 pandemic

The data on the incidence of the COVID-19 pandemic were taken from the Instituto de Salud Carlos III.<sup>4</sup> We used two variables: number of detected cases (positive diagnosis of active infection) and the number of deaths.<sup>5</sup> For both variables, the data were computed over 100,000 inhabitants. Although data on



hospitalised cases and emergency room admissions were also available, we did not use that data in this analysis because the results were very similar to those offered by the two variables finally included. Detected cases consider the number of reported cases confirmed with a positive diagnostic test for active infection (PDIA) as established in the strategy for early detection, surveillance, and control of COVID-19 and the cases notified before 11 May 2020 that at any time required hospitalisation, admission to the ER, or resulted in death with a clinical diagnosis of COVID-19 according to the current case definitions.

The dates that were considered for the detected cases included the date of diagnosis and, in its absence, the date of declaration to the community (see Appendix S1 for a more detailed explanation of the key date). For deaths, the date of death was considered. The data were gathered for each wave, using the following schedule:

- 1st wave: from 01/03/2020 to 26/06/2020 (peak 26/03)
- 2nd wave: from 27/06/2020 to 10/12/2020 (peak 4/11)
- 3rd wave: from 11/12/2020 to 16/03/2021 (peak 27/01)
- 4th wave: from 17/03/2021 to 22/06/2021 (peak 26/04)
- 5th wave: from 23/06/2021 to 14/10/2021 (peak 27/07)
- 6th wave: from 15/10/2021 to 10/03/2022 (peak 21/01)

All the information is provided at the provincial level for Spain, spanning from February 2020 to March 2022. Because the restrictions and health policies related to the pandemic evolved along the waves, we briefly recall them.

The first case of COVID-19 in Spain occurred on 31 January 2020 and involved a German tourist who was on vacation in the Canary Islands. After that initial case, more began to appear, leading to the declaration of the start of the pandemic in Spain on 1 March 2020 (the beginning of what we define as the first wave). A few days later, the state of emergency was declared (14 March 2020), and the population was placed under lockdown. In mid-April, discussions about the de-escalation began, initiating a process in which time slots were defined for the outing of children, the elderly, and so on. On 21 June, what was referred to as the “new normality” began, in which restrictions on movement throughout Spain were lifted, although a minimum distance of one and a half metres between individuals and mandatory mask usage were required. The second wave was marked by a new increase in infections starting in September and the imposition of a new nationwide state of alarm at the end of October, which

included nighttime curfews and regional border closures. During the third wave, the state of emergency was still in place, but vaccination campaigns were gradually initiated. The fourth wave was characterised by the end of the second state of alarm in early May 2021, when 10% of the country’s population had been vaccinated. The fifth wave progressed much like the previous one, without restrictions except for the use of a mask in certain places (hospitals and medical centres and public transportation), reaching 70% of the population vaccinated in September 2021 (even with three doses for immunocompromised patients and those older than 70). The last wave was characterised by an increase in both the number of cases and deaths, primarily due to the appearance of a new virus variant (Omicron), as well as the continued use of masks as a protective measure.

Finally, Spain was unlike other countries with a more homogeneous territorial management of the pandemic. Following a period of recentralisation of the state-level health policy in terms of pandemic management during the initial state of alarm in Spain, a process of co-governance with the autonomous communities for the transition to a new normality began in May 2020. This process led to regional governments adopting different responses in terms of the imposed restrictions (as exemplified by the contrast between greater leniency in the case of the autonomous community of Madrid compared to more restrictive measures in the region of Catalonia).<sup>6</sup>

### 3.1.2 | Determinants of the COVID-19 pandemic

Based on our analysis of the literature, we chose the variables proxying for the five dimensions of factors that could be relevant in explaining the incidence of COVID-19.

For **climatology**, we selected the variable of average temperature (*Temperature*)<sup>7</sup> taken from the Agri-Cast Resources Portal of the Joint Research Centre of the European Commission.<sup>8</sup> To proxy for the **demographic structure** of the province, we considered the share of people over 70 years old (*Pop\_70*), while population density, *Pop\_dens*, was considered for the existence of **agglomerations** of population (both variables from National Institute of Statistics). Gross domestic product per capita (*GDPpc*) was chosen to proxy the **economic wealth** of each province (taken from INE). Finally, the dimension associated with **connectivity** was captured from three different perspectives: intraurban mobility, passengers’ mobility and freight mobility. In the first case, we used a variable that considers whether the province has a mass transit system (*Subway*) in any of the cities in the province. In the second case, the number of travellers staying in hotels was

selected (*Travellers*), while in the last case, we used information about international trade (exports + imports) over GDP (*Comin\_gdp*), extracted from the Institute of Foreign Trade.

Table S1 in the Appendix online presents the definition, frequency and source of each variable. Although many variables vary across the different waves, the ones referring to the GDP and population structure are maintained throughout the analysis.<sup>9</sup> Appendix S1 also offers the maps of the spatial distribution of the variables selected as potential determinants.

### 3.2 | Description of the COVID-19 incidence in Spanish provinces

Here, our first focus is on the analysis of the evolution of the incidence of the COVID-19 pandemic (case and death rates; that is, number of cases and deaths per 100,000 inhabitants, respectively) across the six waves in Spain. Temporarily speaking (see Tables 1 and 2), there are some substantial differences along the waves.

The first wave presents an interestingly low case rate, probably because of limited ability to detect the pathogen. Afterwards, incidence grew between waves except in the fourth wave, which presents an intriguingly low incidence that corroborates the effectiveness of vaccines as well as the restrictions still imposed. Finally, during the last wave, the highest peak ever reached throughout the entire pandemic in Spain was observed, probably due to more relaxed restrictions and Christmas celebrations and due to the appearance of an Omicron virus variant.

In relation to the death rate, the greatest value occurred during the first wave, although a high magnitude was also observed in the second and third waves. Likewise, from the fourth wave, a very pronounced change in the level was observed, with a lower number of deaths throughout the period (going from a value of 69.83 deaths per 100,000 inhabitants in the first wave to a value of 11.25 in the fourth). Similarly to the number of cases, deaths also grew in the fifth and, especially, in the sixth wave, although the impact of the

vaccines significantly reduced the magnitude of the problem (if we take into consideration the number of deaths in relation to the number of cases).

## 4 | EXPLORATORY SPATIAL ANALYSIS OF COVID-19 IN SPAIN

We turn now to analyse the regional distribution of the incidence of COVID-19 in the provinces in Spain for the six consecutive waves. For the case rate (Figure 1), in general terms, we observe that the pattern of the spatial distribution of the illness had important changes across time. The correlation matrix across the case rate in the six waves (Figure 2) reinforces the conclusion of the absence of high correlations in the incidence across waves, with the third wave as the one with the least similarity compared to the rest.

In looking at the quantile maps of the death rate (Figure 3) as well as the correlation across waves (Figure 4), there seems to be a lower homogeneity between neighbouring provinces than for the case rate, which could be because the contagion of the disease is due to a process of transmission between human beings (and, therefore, their proximity is decisive). Second, the correlation across waves is higher than those observed in terms of the number of cases, indicating more similarities in the spatial distribution between waves in deaths than in cases.

All in all, the territorial patterns for both case and death rates point to a potential presence of a positive spatial autocorrelation process among different provinces<sup>10</sup>: In all six waves of COVID-19, provinces with

TABLE 2 Temporal evolution of COVID-19 in Spain. Death rate.

Variable	Obs	Mean	Min	Max
First wave death	50	69.83	4.07	261.27
Second wave death	50	55.00	8.84	147.72
Third wave death	50	58.90	11.40	145.13
Fourth wave death	50	11.25	1.84	27.98
Fifth wave death	50	13.75	4.94	37.71
Sixth wave death	50	34.41	10.84	81.15

TABLE 1 Temporal evolution of COVID-19 in Spain. Case rate.

Variable	Obs	Mean	Min	Max
First wave incidence	50	564.822	69.492	1722.071
Second wave incidence	50	3241.663	854.351	5539.773
Third wave incidence	50	3025.169	814.921	5549.373
Fourth wave incidence	50	1094.762	231.113	2408.844
Fifth wave incidence	50	2395.46	1490.866	4015.11
Sixth wave incidence	50	13,685.67	5448.053	21,489.65

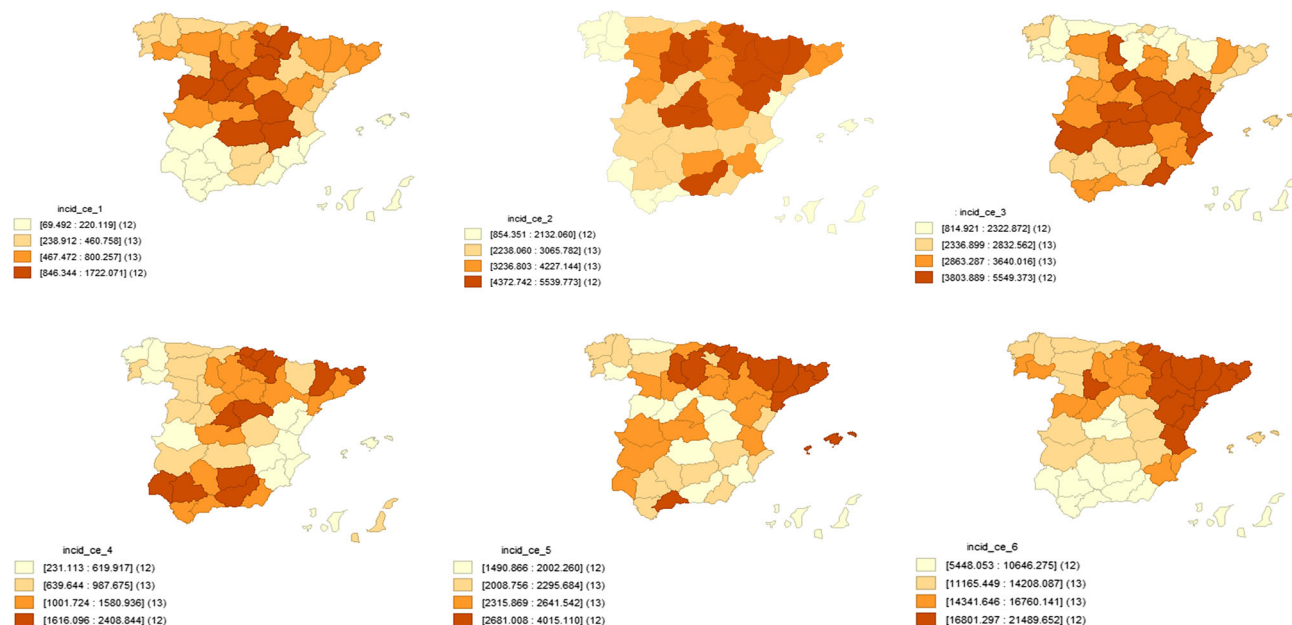


FIGURE 1 Maps of the case rate of COVID-19 in Spanish provinces by waves.

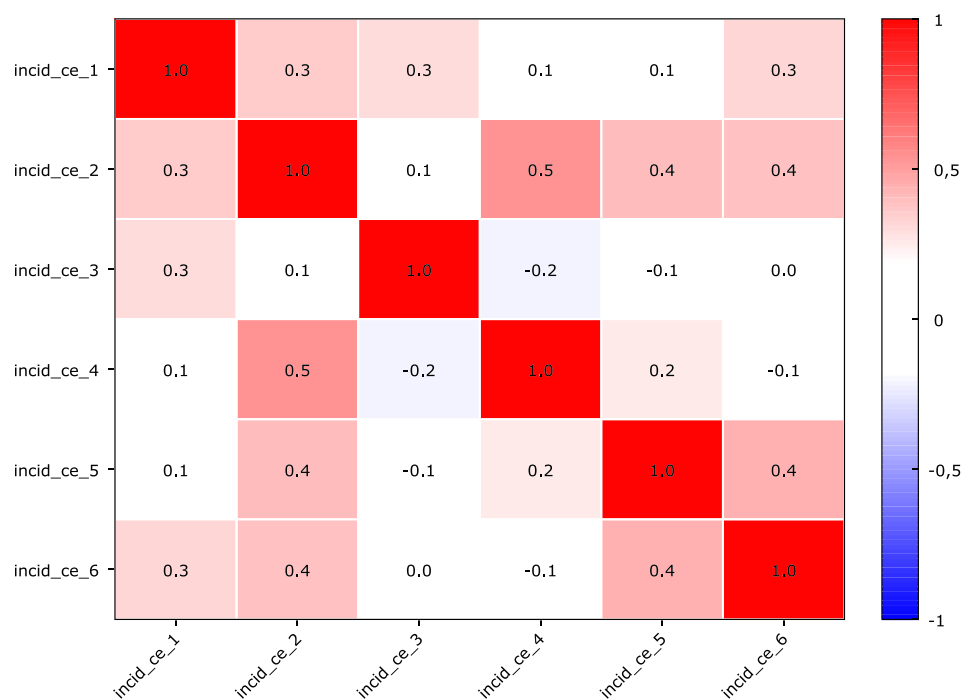


FIGURE 2 Correlation matrix among case rates by waves.

high COVID incidence were surrounded by provinces with high incidence, while provinces with low incidence were surrounded by provinces with low incidence. This pattern was tested with the global Moran's  $I$  test (Tables 3 and 4), with the use of two weight matrices: one based on the first-order physical contiguity criterion and the other based on the inverse of the distance that separates the centroids of each province.<sup>11</sup>

Starting with the case rate, we reject the null hypothesis of no spatial autocorrelation in all the

waves, and therefore confirm that, as shown in the maps, the COVID-19 pandemic followed a clear pattern of positive spatial autocorrelation (in line with the one obtained by González et al., 2022, for Spanish regions and Páez et al., 2021, for Spanish provinces). Additionally, the autocorrelation seems to be more intense when the concept of neighbourhood is extended and the inverse distance matrix is used instead of the first-order physical contiguity matrix, a fact that reinforces the amplitude of the diffusion.

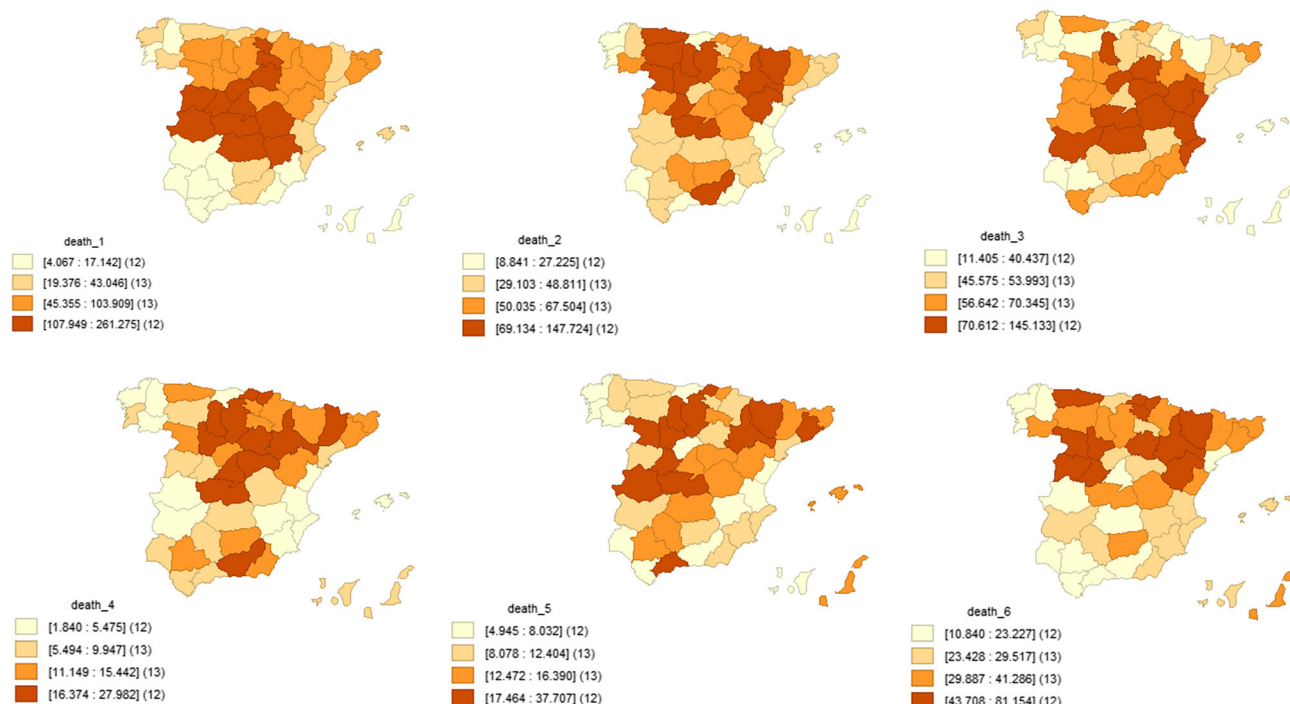
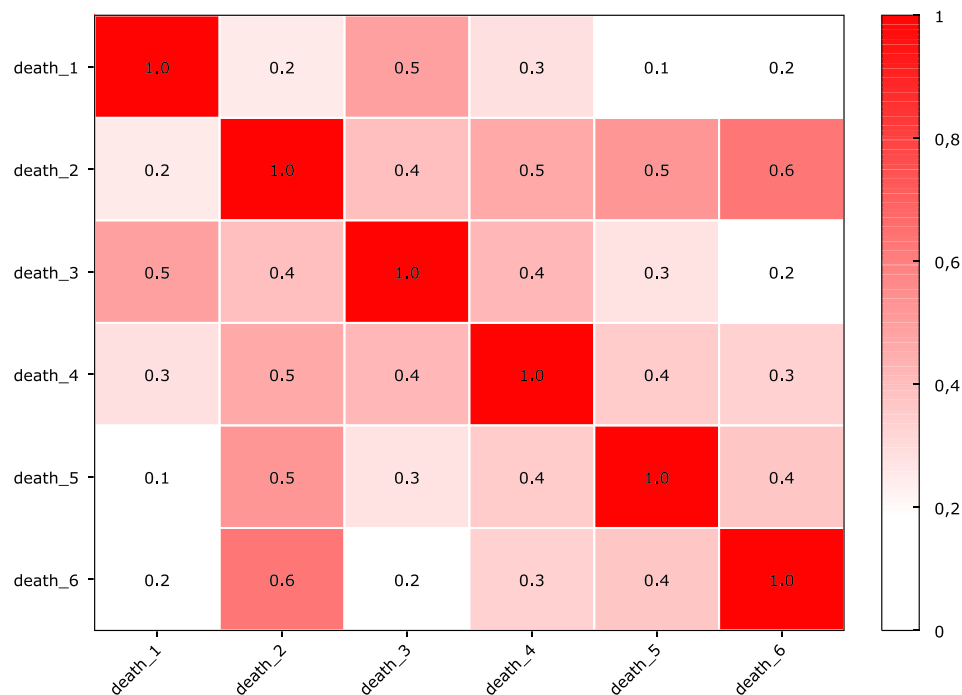


FIGURE 3 Maps of the death rate due to COVID-19 in Spanish provinces by waves.

FIGURE 4 Correlation matrix among death rates by waves.



Likewise, the autocorrelation is more intense in the last wave (where the case rate is also higher) in which there seems to be a division of the country into two parts: a high number of cases in the northeast and fewer cases in the southwest.

When the analysis was repeated for the death rate, the null hypothesis of no spatial autocorrelation is

rejected in practically all the scenarios (apart from the fifth wave) but with a test value lower than for the case rate. Likewise, while the autocorrelation was much stronger in the last wave for the case rate, the greatest spatial association for deaths was observed in the first wave.

Despite its significance, the Global Moran's  $I$  test has some limitations because it only reveals a global



spatial behavioural pattern, even though a variety of local spatial patterns can arise. To check for this limitation, we also performed a local wave-by-wave analysis (Figures 5 and 6). The local indicators of spatial autocorrelation were applied using a weight matrix based on the inverse of the distance between the centroids of each pair of provinces, which presents the advantage

of allowing the analysis to take the islands into consideration.

During the first and second waves, centric and northern provinces represented clusters with the highest values of COVID-19 incidence, whereas southern provinces were low-value clusters. In the third wave, the clusters with the highest incidence were in the centre-eastern provinces, and only in the north in the fourth wave, which changed to the north-eastern area in the fifth and sixth waves, with southern provinces as well as Galicia (northwestern region) being clusters of low cases. In the sixth wave, the number of clusters reached its highest point, totalling 30, with only one spatial outlier. Therefore, in the latest wave, the spatial heterogeneity was quite pronounced and in line with the highest value of the Moran's  $I$  test. All in all, COVID-19 expanded unevenly between Spanish provinces over the six observed waves. Similar conclusions were obtained in the case of the death rate (Figure 6), although the number of provinces that did not end up showing a significant autocorrelation scheme is greater than in the case rate, a result that is also in line with the lower intensity of autocorrelation at a global level obtained with the Moran's  $I$  test.<sup>12</sup>

TABLE 3 Moran's  $I$  test (standardised). Case rate.

Z-Moran ( $p$ value)	Binary contiguity	Inverse distance
First wave incidence	5.22 (0.002)	5.20 (0.002)
Second wave incidence	5.04 (0.002)	6.02 (0.002)
Third wave incidence	4.17 (0.002)	5.55 (0.002)
Fourth wave incidence	5.19 (0.002)	3.57 (0.004)
Fifth wave incidence	4.63 (0.002)	5.43 (0.002)
Sixth wave incidence	8.74 (0.002)	10.11 (0.002)

Note:  $p$  value in brackets.

TABLE 4 Moran's  $I$  test (standardised). Death rate.

Z-Moran ( $p$ value)	Binary contiguity	Inverse distance
First wave death	5.06 (0.002)	4.40 (0.008)
Second wave death	2.60 (0.016)	3.09 (0.01)
Third wave death	1.80 (0.052)	3.45 (0.008)
Fourth wave death	3.03 (0.008)	3.59 (0.006)
Fifth wave death	1.54 (0.064)	0.97 (0.150)
Sixth wave death	3.44 (0.006)	3.57 (0.004)

Note:  $p$  value in brackets.

## 5 | DETERMINANTS OF COVID-19 IN SPANISH PROVINCES: METHOD AND MAIN FINDINGS

We turn now to estimate a model in which the dependent variable is one of the proxies for the incidence

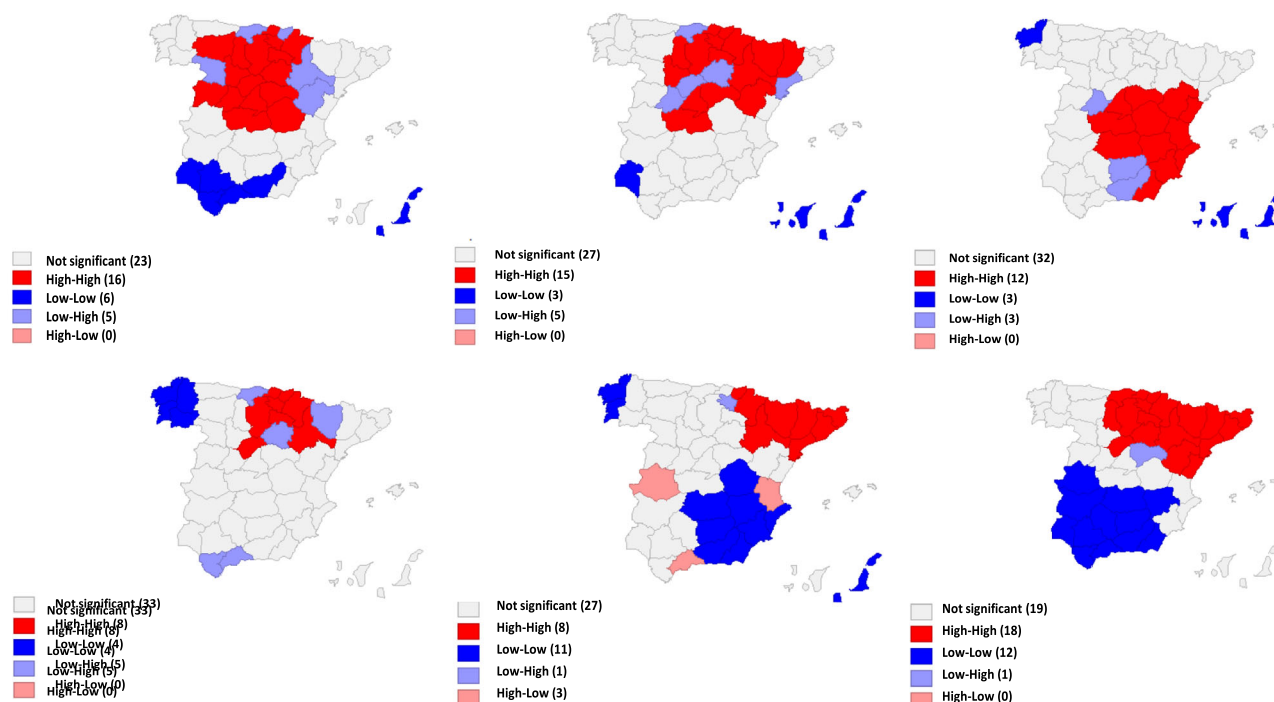


FIGURE 5 Local Moran's  $I$ . Case rate by waves. Inverse distance matrix.

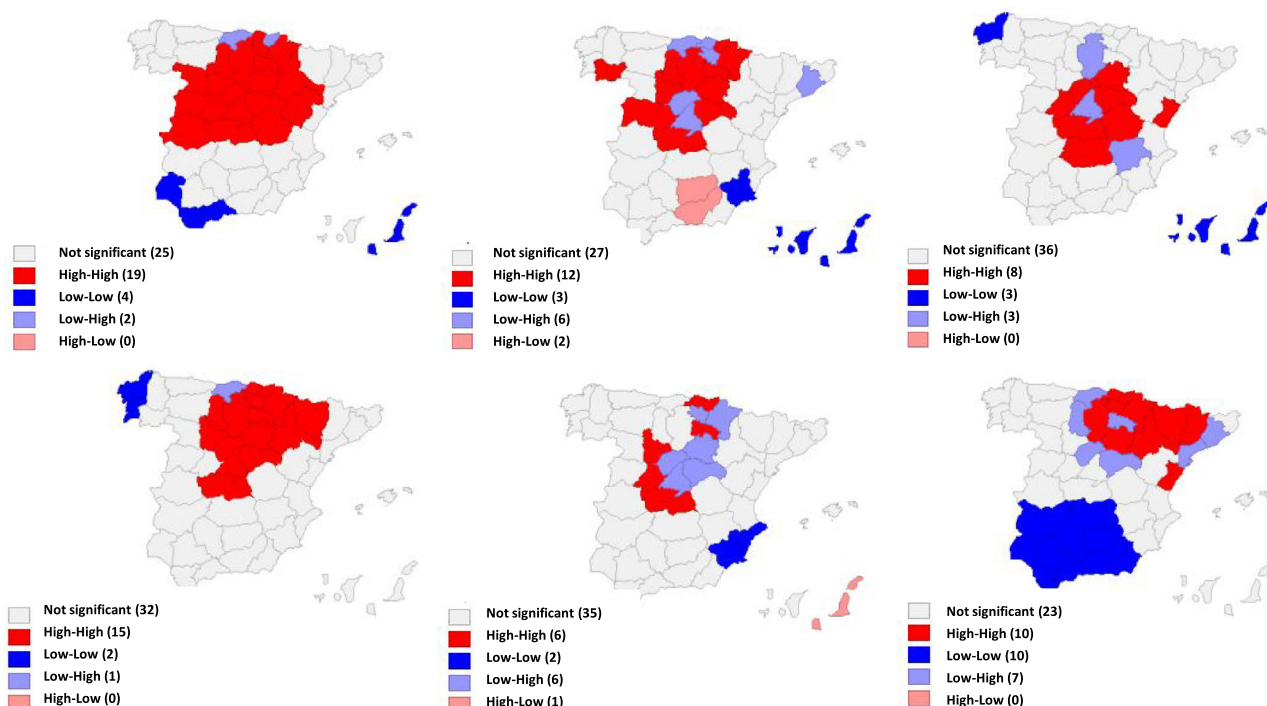


FIGURE 6 Local Moran's  $I$ . Death rate by waves. Inverse distance matrix.

(case and death rates) and the explanatory variables are their potential determinants as defined in Section 4. In addition, a spatial lag of the dependent variable is included ( $WIncidence$ ), which was computed using a spatial weight matrix based on the inverse of the distance between centroids. The decision to include a spatial spillover proxied by the lag of the dependent variable in the proposed initial model has both theoretical and empirical justifications. Theoretically, in the case of COVID-19 transmission, even though the initial spatial distribution of the illness incidence might be random, the virus outbreak would result in the creation of a cluster consisting of the province of origin of the virus and its neighbouring areas. Empirically, the results of the spatial exploratory analysis presented in Section 4 allow us to conclude in favour of a spatial dependency pattern in almost all waves for both variables (case rate and death rate) in the case of the Spanish provinces. Consequently, we estimate the following model:

$$\begin{aligned} Incidence_i = & \beta_0 + \beta_1 Temperature_i + \beta_2 Pob_{70}_i \\ & + \beta_3 Pob_{dens}_i + \beta_4 GDPpc_i + \beta_5 Subway_i \\ & + \beta_6 Travellers_i + \beta_5 Comin_{gdp}_i \\ & + \rho WIncidence_i + u_i \end{aligned}$$

where  $i$  refers to the Spanish provinces in one wave, and we thus run six different regressions, one for each wave. This spatial autoregressive lag model is estimated using a maximum likelihood method that takes explicit account of the endogeneity problem generated

by the inclusion of the spatial lag of the dependent variable (Anselin, 1988).

Tables 5 and 6 present the results of the estimation for the case and death rates, respectively. In all waves, the results are presented both including and excluding the temperature variable because depending on the waves, its inclusion diminishes the significance of other variables (which is especially evident in the case of the first wave). From both tables, the following general conclusions can be drawn. First, the selected variables have a greater capacity to explain the case rate than the death rate (greater model fit for the former). Second, when comparing the goodness of fit between waves, the fourth wave shows the worst fit for the case and death rates,<sup>13</sup> with the fifth wave also having a bad fit for the death rate. Conversely, for both variables, the best fit occurs in the last wave. All the selected explanatory variables are relevant at least in one estimation.

The spatial lag of the case rate (Table 5) is consistently significant even at the 1% level in most waves (the significance of spatial lag was also found by González et al., 2022, for Spanish regions; Páez et al., 2021, for Spanish provinces; Baena-Díez et al., 2020, for districts in the city of Barcelona; and by Maza and Hierro [2022] for municipalities in Madrid). Its highest magnitude is observed in the sixth wave, which is when the highest spatial heterogeneity in the distribution of the case rate is observed, with the highest value of the univariate Moran's  $I$  statistic and the highest number of detected spatial clusters, a total of 30 provinces.

TABLE 5 Regression of case rate of COVID-19 over main determinants.

	(1) 1st wave	(2) 1st wave	(3) 2nd wave	(4) 2nd wave	(5) 3rd wave	(6) 3rd wave	(7) 4th wave	(8) 4th wave	(9) 5th wave	(10) 5th wave	(11) 6th wave	(12) 6th wave
W_lag	0.578***	0.338*	0.645***	0.632***	0.667***	0.618***	0.617***	0.605***	0.345*	0.343*	0.745***	0.752***
Pop_70	0.012	-0.038	-0.024*	-0.036*	-0.017	-0.032**	-0.057**	-0.087***	-7.6E-04	0.001	0.011	0.012
Pop_dens	-0.275**	-0.089	-0.121*	-0.115*	-0.046	0.086	0.083	0.086	-0.095***	-0.092r**	0.012	0.006
GDPpc	1.541***	0.734	0.575**	0.411	-0.295	-0.670**	0.450	-0.021	0.462***	0.483***	0.471***	0.486***
Subway	0.053*	0.039	0.019	0.021	0.034**	0.021	0.027	0.027	0.0167**	0.016*	-0.017**	-0.017*
Travellers	-0.047	-0.052	-0.117	-0.109	-0.109	-0.128	-0.289**	-0.219	0.092**	0.092**	-0.0008	-0.001
Comin_gdp	-0.178	-0.287**	0.010	-0.007	0.061	0.018	0.079	0.040	0.034	0.036	0.018	0.020
Temperature		-3.501***		-0.615		-0.606***		-1.419*		0.089		0.042
Constant	-11.114**	7.978	-0.718	2.977	7.077**	12.509***	1.944	10.33*	-0.486	-0.991	-2.575*	-2.888
N	50	50	50	50	50	50	50	50	50	50	50	50
Akaike	94.47	85.79	32.01	33.07	30.04	25.43	69.69	68.33	-36.89	-34.97	-19.04	-17.09
R <sup>2</sup>	0.62	0.68	0.61	0.61	0.53	0.59	0.40	0.44	0.53	0.53	0.75	0.75

Note: p value of individual significance test.  
\*p < 0.10, \*\*p < 0.05, and \*\*\*p < 0.01.

TABLE 6 Regression of death rates of COVID-19 over main determinants.

	(1) 1st wave	(2) 1st wave	(3) 2nd wave	(4) 2nd wave	(5) 3rd wave	(6) 3rd wave	(7) 4th wave	(8) 4th wave	(9) 5th wave	(10) 5th wave	(11) 6th wave	(12) 6th wave
W_lag	0.559***	0.400**	0.439**	0.421**	0.581***	0.508***	0.456**	0.400**	-0.260	-0.259	0.285	0.238
Pop_70	0.016	-0.041	0.019	0.006	0.003	-0.014	-0.056*	-0.095**	-0.009	0.017	0.006	0.002
Pop_dens	-0.378**	-0.190	-0.255**	-0.248**	-0.108	0.033	0.028	0.031	-0.161*	-0.129	-0.075	-0.051
GDPpc	1.374**	0.485	0.476	0.286	-0.287	-0.687**	0.577	-0.011	0.949**	1.185***	1.060***	0.995**
Subway	0.069*	0.059	0.033	0.036	0.026	0.013	0.034	0.034	0.049**	0.039*	-0.009	-0.013
Travellers	-0.024	-0.036	-0.084	-0.075	-0.072	-0.096	-0.384*	-0.302	-0.084	-0.095	-0.029	-0.026
Comin_gdp	-0.108	-0.228	-0.020	-0.038	0.115	0.075	0.0748	0.024	-0.099	-0.076	-0.163*	-0.170*
Temperature		-3.481***		-0.706		-0.667**		-1.842*		1.066		-0.187
Constant	-10.42	8.799	-0.934	3.245	5.480*	11.214***	0.728	11.499	-4.443	-10.567	-7.072**	-5.882
N	50	50	50	50	50	50	50	50	50	50	50	50
Akaike	120.26	115.89	75.63	77.14	53.21	50.50	105.09	104.46	68.30	68.90	52.41	54.23
R <sup>2</sup>	0.56	0.60	0.53	0.53	0.50	0.54	0.27	0.30	0.25	0.27	0.75	0.75

Note: p value of individual significance test.  
\*p < 0.10, \*\*p < 0.05, and \*\*\*p < 0.01.

The spatial lag of the death rate (Table 6) is also significant but only in the first four waves. This would be in line with the result obtained from the univariate analysis in Section 4 that detected spatial randomness in the distribution of the number of deaths in the fifth wave (when the number of significant spatial clusters equalled the number of spatial outliers). However, in the sixth wave, global univariate spatial dependence was detected; thus, the reason for the nonsignificance of the spatial lag coefficient in this wave could be the inclusion of the *GDPpc* variable, which shows a spatial distribution almost identical to that of the death rate.<sup>14</sup> If compared with the case rate model, the estimated parameter of the spatial lag in the case of the death rate is also lower in magnitude. That is, spatial dependence seems to have a greater role in terms of incidence transmission than for deaths (a result consistent with the conclusions of the univariate analysis in which the case rates exhibited higher Moran's *I* values compared to those observed for the death rates).

As for the determinants of the incidence of COVID-19, the *Temperature* variable is relevant in the first, third, and fourth waves, both for cases and deaths, although it is in the first wave where the coefficient's magnitude and significance are greater, always with a negative sign in line with the literature (for example, Páez et al., 2021). The inclusion of the *Temperature* variable (when relevant) reduces the magnitude of the spatial lag, showing that the spatial pattern of the dependent variable is partly picked up by the spatial pattern of the *Temperature* variable. Additionally, the inclusion of this variable in the first wave (when the coefficient associated with the variable is the highest) absorbs most of the explanatory capacity of the other determinants, because *POp\_dens*, *GDPpc*, or *Subway* lose their significance when *Temperature* is included. We thus conclude that the *Temperature* variable (relevant in epidemiological studies) may have contributed to a fast spread of the virus during the first wave in which the spatial pattern of incidence and death variables could be the result of several months of virus transmission (before the state of alarm and, therefore, without protective measures).

The *Percentage of population over 70 years old* does not seem to have been particularly relevant for any of the two variables proxying for the incidence of the illness, except in the case of the fourth wave. In such a case, the negative sign of the coefficient is in line with Páez et al. (2021) for Spanish provinces but contrasts with the positive sign of the “average age” variable used by Maza and Hierro (2022) for the municipalities in Madrid. In this fourth wave, it was almost the only variable (in addition to the spatial lag of the dependent variable) that proved to be relevant. This result could be explained by the fact that the vaccination process was more extensive in the fourth wave, with 10%

of the Spanish population having been vaccinated in May. Because vaccination priority was mainly based on age, this could explain both the relevance and the sign of this variable (a higher share of the population over 70, who were more vulnerable, would result in a larger population rate protected by vaccines and corresponding fewer cases and deaths).

The *Pop\_dens* variable was especially relevant and negative in the fifth wave for the case rate and in the second wave for the death rate. The negative sign goes against the expectation that contact rates are higher in more dense areas and, consequently, positively correlated with the transmission of the virus (as detected by González et al., 2022, for Spanish regions or Maza and Hierro [2022] for municipalities in Madrid). A potential explanation is in line with what was obtained in other works, such as Páez et al. (2021), which was also for Spanish provinces, in which they argued that this negative correlation is due to the so-called risk compensation, that is, a situation in which people adapt their behaviour according to the perceived level of risk and become more careful when the perceived risk is higher, as in dense areas, and vice versa.

The level of *GDPpc* in each province is mainly relevant and positive when explaining both incidence rate indicators, with special relevance in the fifth and sixth waves (with the only exception of a negative sign in the third wave). Especially in the sixth wave, the spatial distribution of *GDPpc* is very similar to the spatial distribution of the incidence rates. The positive sign of this variable contradicts the negative sign obtained by Baena-Díez et al. (2020) in the first wave for the districts of Barcelona or by Maza and Hierro (2022) for municipalities in Madrid (that is, very small geographic scales) but is consistent with the positive sign obtained for this same variable for larger areas, such as Spanish provinces in Páez et al. (2021) or with the negative sign of the unemployment rate variable for Spanish regions detected by González et al. (2022). This positive relationship would indicate that wealthier provinces may remain more active, in relative terms, even during a lockdown due to the presence of more non-traded activities. Another explanation relates to the fact that wealthier provinces can be part of global city networks with a greater “potential to be further ahead in the trajectory of the pandemic” (Páez et al., 2021, p. 399). In addition, the higher relevance of this variable in the last two waves suggests that during a period of greater vaccination and normality (fewer restrictions), one main reason for a higher number of cases and deaths could be a greater economic activity (resulting in increased internal movement of people, higher commuting, and increased activity in certain service sectors).

As for the role of connectivity on the diffusion of the virus, the *Subway* variable has proven to be relevant only in some waves, particularly the fifth one, for both dependent variables. The sign is generally positive,



which is in line with what was found by Páez et al. (2021), inferring that thanks to higher vaccination rates (with 70% of the population vaccinated in October 2021), the higher internal mobility in public transportation could explain a higher transmission and risk of death. This result would point to public transportation as potential hotspots for social contact. As for the across-provinces connectivity, the *Travellers* variable is not particularly significant in any wave, except again in the fifth wave for the number of cases and with a positive sign, in line with what was found by Maza and Hierro (2022). Finally, the *Comin-GDP* variable proxying for freight mobility is the least relevant variable, being significant and negative only in the first wave for the case rate and in the sixth wave for the death rate. The provinces with the main ports for international goods—Cádiz, Valencia, Barcelona, Murcia, Huelva, Bilbao, Tarragona and Las Palmas—and the border provinces were the ones that showed lower incidences in the first wave. That is because the majority of cases were concentrated in the central part of the country, thereby providing an explanation for the negative sign.

## 6 | DISCUSSION

This paper has been focused on the incidence of COVID-19 in the Spanish provinces along the six waves of the pandemic and has two main objectives. First, we compared the uneven spatial distribution of the incidence, both in terms of the number of cases and the number of deaths per 100,000 inhabitants, across the six waves considering exploratory spatial techniques. Second, we studied the main determinants of the diffusion of the pandemic through the estimation of spatial models, while analysing to what extent they experimented a different role across waves.

As for the spatial distribution of the pandemic, we obtained a clear generalised pattern of positive spatial autocorrelation between nearby provinces, both in terms of cases and deaths, although their intensity was higher in the former. This points to the incidence of the COVID-19 as an interprovincial contagion process. At a local level, it seems that COVID-19 expanded unevenly between Spanish provinces over the six observed waves. In particular, the southern provinces appear as clusters with low values as far as the number of cases is concerned, while the opposite occurs with the central provinces in the first wave and the northeastern provinces in the last. We also observe that the spatial association found in the distribution of the incidence of the illness is more intense when the concept of neighbourhood is extended, reinforcing the extent of the diffusion.

As for the evolution in time, we observe that the pattern of the spatial distribution of the case and death rates revealed important changes along time, with the correlation across waves being higher in terms of

deaths. As for the determinants of the incidence of the pandemic, at its onset, when the first state of emergency was declared after a previous incubation period of the virus in the preceding two months, a factor such as temperature—which is significant in epidemiological studies—was highly relevant. However, after the implementation of mobility restrictions and the widespread vaccination process, variables such as GDP became more significant. This explains how increased economic activity (associated with lower unemployment rates and greater population mobility) could lead to a higher number of cases and deaths. In addition, during the initial stages of the vaccination process when specific groups, such as the elderly (over 70), were prioritised, provinces with a higher rate of these specific groups had a lower number of infections and deaths because a larger proportion of vulnerable individuals had been vaccinated. Although some of these findings are in line with previous evidence for Spain from other authors, this is especially the case in those papers using regional or province data and not for other, more disaggregated levels of analysis. It seems, therefore, that the geographical scale used is important, and general statements cannot be made without considering the level of geographical disaggregation used. Because Spain's settlement pattern resembles many other southern European countries such as Italy, France, Portugal, and Greece, our findings could be expanded to other settings beyond Spain in a generalisable way.

Our findings on the evolution of COVID-19 are deeply intertwined with the theory of diffusion, which seeks to understand how infectious diseases spread through populations and across geographic territories. Initially, during the onset of the pandemic, transmission patterns were influenced by factors commonly associated with infectious diseases, such as environmental conditions. However, the spatial patterns of COVID-19 incidence were far from static; they exhibited a remarkable degree of dynamism over time. As governments and public health authorities implemented interventions to curb transmission, such as lockdowns, social distancing measures, and mask mandates, the spatial dynamics of the pandemic began to shift. Areas that initially experienced lower incidence rates may have become new hotspots as the virus found opportunities to propagate within vulnerable populations or through social networks. That is, mobility restrictions, social distancing measures, and mask mandates disrupted traditional modes of disease transmission, leading to a reconfiguration of spatial patterns as transmission rates fluctuated in response to varying levels of compliance and enforcement. The emergence of new variants of the virus further complicated the spatial dynamics of COVID-19 transmission. Variants with increased transmissibility could lead to shifts in incidence patterns. The dynamic interplay between viral evolution

and human behaviour underscored the need for adaptive public health responses that could quickly identify and contain emerging hotspots of transmission.

## 7 | CONCLUSION

All in all, we can conclude that we cannot refer to a single pandemic, but rather many pandemics in waves. Significant differences in the spatial distribution of the number of cases and deaths across the six waves demonstrate that each wave has unique characteristics. The low correlation between waves, especially in the number of cases, and the changes in their determining factors among the different waves further reinforce this idea.

Similarly, just as we cannot refer to a single pandemic, we cannot draw homogeneous conclusions at the national level. The notable regional differences detected in the spatial distribution of the pandemic, which conceal behind them interregional variations in climatological issues, economic dynamism, sectoral specialisation, and socio-health resources, suggest the need for a subnational scale response. Furthermore, the positive and significant spatial autocorrelation detected among neighbouring provinces supports the implementation of responses at the regional level. However, this does not preclude the adoption of specific measures at the local level when necessary.

In addition, the geographic scale used in any analysis of the pandemic is crucial in determining the positive or negative influence of certain factors. For instance, at the provincial level, a positive influence of GDP per capita on the number of COVID-19 cases and deaths has been detected in our study. In contrast, other studies applied to Spain at a more local level (districts within the same city or municipalities) have found the opposite effect. This indicates how the relationship between socioeconomic factors and the spread of the virus depends on the level of spatial disaggregation in the analysis.

Moreover, the main determinants of the incidence of the COVID-19 depend on the stage of the pandemic we are in. Traditional factors of geographic virus spread, such as the temperature of the area, seem to be particularly important at certain stages of a pandemic, especially at its onset. In contrast, at more advanced stages, when the vaccination process is further along, variables such as the economic dynamism of the area may become more relevant. This shift in determining factors throughout the different stages underscores the need for flexible and adaptive approaches in pandemic management, capable of responding to changing conditions and diverse regional contexts.

In conclusion, the evolution of COVID-19 has exemplified the dynamic nature of infectious disease

diffusion, with spatial patterns of incidence shifting in response to a complex interplay of factors, including human behaviour, public health interventions, viral evolution, and vaccination efforts. Understanding these dynamics is crucial for devising effective strategies to mitigate the spread of the virus and minimise its impact on populations worldwide. While nothing can be done with respect to the meteorological conditions and little can be done about agglomeration characteristics, regional governments can respond to the pandemic with the provision of an efficient vaccination procedure and other measures adapted to the characteristics of the region under their government.

## ACKNOWLEDGEMENTS

We acknowledge financial support from Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR), Generalitat de Catalunya, Project 2020PANDE00060.

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

## DATA AVAILABILITY STATEMENT

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

## ETHICS STATEMENT

This research was conducted in compliance with all relevant ethical guidelines and regulations approved by the Committee on Ethics in Research of the University of Barcelona (CEI-UB).

## ORCID

Rosina Moreno  <https://orcid.org/0000-0002-3500-8622>

## ENDNOTES

- <sup>1</sup> We give this figure in 2022 since it is the end of the period considered in this paper.
- <sup>2</sup> Other papers that have followed the specific purpose of investigating the spatial distribution of the disease through the consideration of spatial techniques have focused on different countries and cities, such as Murugesan et al. (2020) in India, Mota et al. (2021) in Brazil, Guliyev (2020) for China, and Maiti et al. (2021) for the United States, among others.
- <sup>3</sup> In this paper, these other factors were not considered, not only because of the lack of data at the provincial level but also because we do not expect them to have a significant influence since there is not variability across Spanish provinces in the level of institutional quality, social factors, and vaccination levels.
- <sup>4</sup> Instituto de Salud Carlos III: <https://cnecovid.isciii.es/covid19/#documentaci%C3%B3n-y-datos> (last access 24 November 2022).
- <sup>5</sup> As pointed out by Karlinsky and Kobak (2021) and Mathieu et al. (2020), comparing the impact of the COVID-19 pandemic among countries or across time is difficult since the reported number of cases and deaths can be strongly affected by testing capacity and reporting policy. Instead of deaths, excess mortality (increase in

all-cause mortality relative to the expected mortality) is widely considered as a more objective indicator of COVID-19 deaths (Bartoszek et al., 2020). However, data on excess mortality is not available at the provincial level in Spain (we only found the information at the national level). In any case, we would expect that the main determinants of the number of deaths would not be highly different from the determinants of excess mortality across provinces in the same country, given that the policies to control the pandemics were centralised.

- <sup>6</sup> While regional discretion in pandemic management may have influenced its varied evolution, according to Biglino Campos (2021), it has not been proven that the territorial structure or the distribution of competencies corresponding to each form of state has been a determining factor in the success or failure in the fight against the pandemic. Thus, the results do not seem to have differed much between unitary states, such as France, and federal states, such as the Federal Republic of Germany.
- <sup>7</sup> Additionally, we selected the variable “rainfall”, although it was not ultimately chosen due to showing lower correlation with COVID-19 incidence compared to the temperature variable.
- <sup>8</sup> <https://agri4cast.jrc.ec.europa.eu/dataportal/> (last access 24 November 2022).
- <sup>9</sup> The variables referring to the GDPpc and population structure remain constant throughout the six waves. There were different reasons behind this procedure. First, because at the time of conducting the research, there was not sufficiently updated information available for the studied period at a provincial level. This led us to keep the variables constant throughout the six waves. Second, because the waves refer to periods of time that do not cover the whole year but only some months and they refer to different years, whereas the data refer to the whole year. Consequently, it was not possible to get the specific data of such variables for each wave. Third, these variables were not expected to present substantial changes in their spatial distribution along the 2 years under consideration; thus, we expect that the influence this may have had should be minimal.
- <sup>10</sup> For a detailed explanation of spatial dependence and its treatment in a regression model, see Anselin (1988), Fotheringham (2009), or LeSage and Pace (2009).
- <sup>11</sup> Spain is very well connected by air, land and sea, a fact that would justify the suitability of defining an alternative weight matrix to physical contiguity based, for example, on the flows of human mobility between provinces (as suggested by Orea & Álvarez, 2022). However, daily information on such mobility flows between Spanish provinces was not available for the entire period analysed in this paper.
- <sup>12</sup> Figures 5 and 6 are also provided in Appendix S2 when using the binary contiguity matrix.
- <sup>13</sup> The fourth wave has the lowest number of both deaths and cases, with the exception of the first wave, which had a lower number of cases. However, this is not real since the capacity for mass testing and monitoring was considerably weak at that time.
- <sup>14</sup> If GDPpc is not included in the estimation, the spatial lag of the dependent variable in Table 6 becomes significant.

## REFERENCES

- Ahmadi, M., Sharifi, A., Dorosti, S., Jafarzadeh Ghouschi, S., & Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Science of the Total Environment*, 729, 138705. <https://doi.org/10.1016/j.scitotenv.2020.138705>
- Ahmed, S. F., Quadeer, A. A., & McKay, M. R. (2020). Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-Cov-2) based on SARS-Cov immunological studies. *Viruses*, 12(3), 254. <https://doi.org/10.3390/v12030254>
- Almagro, M., & Orane-Hutchinson, A. (2020). JUE insight: The determinants of the differential exposure to COVID-19 in New York City and their evolution over time. *Journal of Urban Economics*, 127, 103293. <https://doi.org/10.1016/j.jue.2020.103293>
- Amdaoud, M., Arcuri, G., & Levratto, N. (2021). Are regions equal in adversity? A spatial analysis of spread and dynamics of COVID-19 in Europe. *The European Journal of Health Economics*, 22(4), 629–642. <https://doi.org/10.1007/s10198-021-01280-6>
- Anselin, L. (1988). *Spatial econometrics: Methods and models* (Vol. 4). Springer Science & Business Media. <https://doi.org/10.1007/978-94-015-7799-1>
- Baena-Diez, J. M., Barroso, M., Cordeiro-Coelho, S. I., Díaz, J. L., & Grau, M. (2020). Impact of COVID-19 outbreak by income: Hitting hardest the most deprived. *Journal of Public Health*, 42(4), 698–703. <https://doi.org/10.1093/pubmed/fdaa136>
- Bartoszek, K., Guidotti, E., Iacus, S. M., & Okrój, M. (2020). Are official confirmed cases and fatalities counts good enough to study the COVID-19 pandemic dynamics? A critical assessment through the case of Italy. *Nonlinear Dynamics*, 101, 1951–1979. <https://doi.org/10.1007/s11071-020-05761-w>
- Bauer, L., Broady, K., Edelberg, W., & O'Donnell, J. (2020). *Ten facts about COVID-19 and the US economy* (Vol. 17). Brookings Institution.
- Bayode, T., Popoola, A., Akogun, O., Siegmund, A., Magidimisha-Chipungu, H., & Ipingbemi, O. (2022). Spatial variability of COVID-19 and its risk factors in Nigeria: A spatial regression method. *Applied Geography*, 138, 102621. <https://doi.org/10.1016/j.apgeog.2021.102621>
- Bhadra, A., Mukherjee, A., & Sarkar, K. (2021). Impact of population density on Covid-19 infected and mortality rate in India. *Modeling Earth Systems and Environment*, 7(1), 623–629. <https://doi.org/10.1007/s40808-020-00984-7>
- Biglino Campos, P. (2021). El impacto de la covid en la distribución de competencias. Tudela Aranda, J. (coord.) “Estado Autonómico y covid-19”, Colección Obras colectivas, Fundación Manuel Giménez Abad, Zaragoza, <https://doi.org/10.47919/FMGA.OC21.0201>
- Bissell, D. (2021). A changing sense of place: Geography and COVID-19. *Geographical Research*, 59(2), 150–159. <https://doi.org/10.1111/1745-5871.12465>
- Briz-Redón, Á., & Serrano-Aroca, Á. (2020). A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Science of the Total Environment*, 728, 138811. <https://doi.org/10.1016/j.scitotenv.2020.138811>
- Brodeur, A., Gray, D., Islam, A., & Bhuiyan, S. (2021). A literature review of the economics of COVID-19. *Journal of Economic Surveys*, 35(4), 1007–1044. <https://doi.org/10.1111/joes.12423>
- Burton, A. L. (2021). Journaling the COVID-19 pandemic: Locality, scale, and spatialised bodies. *Geographical Research*, 59(2), 217–227. <https://doi.org/10.1111/1745-5871.12459>
- Cao, Y., Hiyoshi, A., & Montgomery, S. (2020). COVID-19 case-fatality rate and demographic and socioeconomic influencers: Worldwide spatial regression analysis based on country-level data. *BMJ Open*, 10(11), e043560. <https://doi.org/10.1136/bmjopen-2020-043560>
- Carozzi, F., Provenzano, S., & Roth, S. (2020). Urban density and COVID-19. Centre for Economic Performance (CEP) discussion paper, n1 1711. Available at SSRN 3643204, Urban Density and Covid-19, <https://doi.org/10.2139/ssrn.3643204>
- Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840), 82–87. <https://doi.org/10.1038/s41586-020-2923-3>
- Coccia, M. (2020). Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral



- infectivity similar to COVID. *Science of the Total Environment*, 729, 138474. <https://doi.org/10.1016/j.scitotenv.2020.138474>
- Coelho, M. T. P., Rodrigues, J. F. M., Medina, A. M., Scalco, P., Terribile, L. C., Vilela, B., & Dobrovolski, R. (2020). Global expansion of COVID-19 pandemic is driven by population size and airport connections. *PeerJ*, 8, e9708. <https://doi.org/10.7717/peerj.9708>
- Dalziel, B. D., Kissler, S., Gog, J. R., Viboud, C., Bjørnstad, O. N., Metcalf, C. J. E., & Grenfell, B. T. (2018). Urbanization and humidity shape the intensity of influenza epidemics in US cities. *Science*, 362(6410), 75–79. <https://doi.org/10.1126/science.aat6030>
- Ehlert, A. (2021). The socio-economic determinants of COVID-19: A spatial analysis of German county level data. *Socio-Economic Planning Sciences*, 78, 101083. <https://doi.org/10.1016/j.seps.2021.101083>
- Elhorst, J. P. (2014). *Spatial econometrics from cross-sectional data to spatial panels*. Springer. <https://doi.org/10.1007/978-3-642-40340-8>
- Elliott, P., & Wartenberg, D. (2004). Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives*, 112(9), 998–1006. <https://doi.org/10.1289/ehp.6735>
- Fatima, M., O'Keefe, K. J., Wei, W., Arshad, S., & Gruebner, O. (2021). Geospatial analysis of COVID-19: A scoping review. *International Journal of Environmental Research and Public Health*, 18(5), 2336. <https://doi.org/10.3390/ijerph18052336>
- Ficetola, G. F., & Rubolini, D. (2021). Containment measures limit environmental effects on COVID-19 early outbreak dynamics. *Science of the Total Environment*, 761, 144432. <https://doi.org/10.1016/j.scitotenv.2020.144432>
- Fonseca-Rodríguez, O., Gustafsson, P. E., San Sebastián, M., & Connolly, A. M. F. (2021). Spatial clustering and contextual factors associated with hospitalisation and deaths due to COVID-19 in Sweden: A geospatial nationwide ecological study. *BMJ Global Health*, 6(7), e006247. <https://doi.org/10.1136/bmjgh-2021-006247>
- Fotheringham, A. S. (2009). "The problem of spatial autocorrelation" and local spatial statistics. *Geographical Analysis*, 41(4), 398–403. <https://doi.org/10.1111/j.1538-4632.2009.00767.x>
- Gong, Y., & Zhao, G. (2022). Wealth, health, and beyond: Is COVID-19 less likely to spread in rich neighborhoods? *PLoS ONE*, 17(5), e0267487. <https://doi.org/10.1371/journal.pone.0267487>
- González, I., Pascual Sáez, M., & Cantero Prieto, D. (2022). The dynamics of COVID-19: An empirical analysis with a view to spatial health econometrics using macrodata. *Working Papers. Collection B: Regional and sectoral economics* 2201, Universidade de Vigo, GEN - Governance and Economics Research Network.
- Guliyev, H. (2020). Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics*, 38, 100443. <https://doi.org/10.1016/j.spasta.2020.100443>
- Gullón, P., Cuesta-Lozano, D., Cuevas-Castillo, C., Fontán-Vela, M., & Franco, M. (2022). *Temporal trends in within-city inequities in COVID-19 incidence rate by area-level deprivation in Madrid* (Vol. 76) (102830). Health & place. <https://doi.org/10.1016/j.healthplace.2022.102830>
- Hierro, M., & Maza, A. (2023). Spatial contagion during the first wave of the COVID-19 pandemic: Some lessons from the case of Madrid, Spain. *Regional Science Policy & Practice*, 15(3), 474–492. <https://doi.org/10.1111/rsp3.12522>
- IMF. (2020). <https://www.imf.org/en/Blogs/Articles/2020/04/14/blog-weo-the-great-lockdown-worst-economic-downturn-since-the-great-depression> (last access 26/01/2023).
- Iyanda, A. E., Boakye, K. A., Lu, Y., & Oppong, J. R. (2022). Racial/ethnic heterogeneity and rural-urban disparity of COVID-19 case fatality ratio in the USA: A negative binomial and GIS-based analysis. *Journal of Racial and Ethnic Health Disparities*, 9(2), 708–721. <https://doi.org/10.1007/s40615-021-01006-7>
- Jamil, T., Alam, I., Gojbori, T., & Duarte, C. M. (2020). No evidence for temperature-dependence of the COVID-19 epidemic. *Frontiers in Public Health*, 8, 436. <https://doi.org/10.3389/fpubh.2020.00436>
- Karlinsky, A., & Kobak, D. (2021). Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *eLife*, 10, e69336. <https://doi.org/10.7554/eLife.69336>
- LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781420064254>
- Liang, Z., Zhu, H., Wang, X., Jing, B., Li, Z., Xia, X., & Sun, B. (2020). Adjuvants for coronavirus vaccines. *Frontiers in Immunology*, 11, 589833. <https://doi.org/10.3389/fimmu.2020.589833>
- Liu, Y., Mao, B., Liang, S., Yang, J. W., Lu, H. W., Chai, Y. H., & Xu, J. F. (2020). Association between age and clinical characteristics and outcomes of COVID-19. *European Respiratory Journal*, 55(5), 2001112. <https://doi.org/10.1183/13993003.01112-2020>
- Ma, S., Lai, X., Chen, Z., Tu, S., & Qin, K. (2020). Clinical characteristics of critically ill patients co-infected with SARS-CoV-2 and the influenza virus in Wuhan, China. *International Journal of Infectious Diseases*, 96, 683–687. <https://doi.org/10.1016/j.ijid.2020.05.068>
- Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., & Luo, B. (2020). Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Science of the Total Environment*, 724, 138226. <https://doi.org/10.1016/j.scitotenv.2020.138226>
- Maiti, A., Zhang, Q., Sannigrahi, S., Pramanik, S., Chakraborti, S., Cerda, A., & Pilla, F. (2021). Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustainable Cities and Society*, 68, 102784. <https://doi.org/10.1016/j.scs.2021.102784>
- Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2020). Coronavirus Pandemic (COVID-19). Published online at [OurWorldInData.org](https://ourworldindata.org). Retrieved from: <https://ourworldindata.org/coronavirus>. [Online Resource].
- Maza, A., & Hierro, M. (2022). Modelling changing patterns in the COVID-19 geographical distribution: Madrid's case. *Geographical Research*, 60(2), 218–231. <https://doi.org/10.1111/1745-5871.12521>
- Mena, G. E., Martínez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O., & Santillana, M. (2021). Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science*, 372(6545), eabg5298. <https://doi.org/10.1126/science.abg5298>
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*, 728, 138884. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Morley, J. E., & Vellas, B. (2020). COVID-19 and older adult. *The Journal of Nutrition, Health & Aging*, 24(4), 364–365. <https://doi.org/10.1007/s12603-020-1349-9>
- Mota, C. R., Bressani-Ribeiro, T., Araújo, J. C., Leal, C. D., Leroy-Freitas, D., Machado, E. C., & Chernicharo, C. A. (2021). Assessing spatial distribution of COVID-19 prevalence in Brazil using decentralised sewage monitoring. *Water Research*, 202, 117388. <https://doi.org/10.1016/j.watres.2021.117388>
- Murugesan, B., Karuppanan, S., Mengistie, A. T., Ranganathan, M., & Gopalakrishnan, G. (2020). Distribution and trend analysis of COVID-19 in India: Geospatial approach. *Journal of Geographical Studies*, 4(1), 1–9. <https://doi.org/10.21523/gcj5.20040101>
- Onder, G., Rezza, G., & Brusaferro, S. (2020). Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *Jama*, 323(18), 1775–1776. <https://doi.org/10.1001/jama.2020.4683>



- Orea, L., & Álvarez, I. C. (2022). How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. *Health Economics*, 31(1), 154–173. <https://doi.org/10.1002/hec.4437>
- Páez, A., Lopez, F., Menezes, T., Cavalcanti, R., & Galdino da Rocha, M. (2021). A spatio-temporal analysis of the environmental correlates of COVID-19 incidence in Spain. *Geographical Analysis*, 53, 397–421. <https://doi.org/10.1111/gean.12241>
- Pequeno, P., Mendel, B., Rosa, C., Bosholn, M., Souza, J. L., Baccaro, F., & Magnusson, W. (2020). Air transportation, population density and temperature predict the spread of COVID-19 in Brazil. *PeerJ*, 8, e9322. <https://doi.org/10.7717/peerj.9322>
- Rodríguez-Pose, A., & Burlina, C. (2021). Institutions and the uneven geography of the first wave of the COVID-19 pandemic. *Journal of Regional Science*, 61(4), 728–752. <https://doi.org/10.1111/jors.12541>
- Romero, L. D., & Arroyo, M. (2022). Geografías de la pandemia COVID-19 en España: el retorno del medio rural. *Documents d'Anàlisi Geogràfica*, 2022, vol. 68, num. 1, p. 139–166.
- Sajadi, M. M., Habibzadeh, P., Vintzileos, A., Shokouhi, S., Miralles-Wilhelm, F., & Amoroso, A. (2020). *Temperature, humidity, and latitude analysis to predict potential spread and seasonality for COVID-19*. Social Science Research Network.
- Shi, P., Dong, Y., Yan, H., Li, X., Zhao, C., Liu, W., & Xi, S. (2020). The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak-evidence from China. *MedRxiv*.
- Statista. (2023). <https://www.statista.com/statistics/1110187/coronavirus-incidence-europe-by-country/> (last access 26/01/2023).
- Sun, Y., Hu, X., & Xie, J. (2021). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *Science of the Total Environment*, 758, 143595. <https://doi.org/10.1016/j.scitotenv.2020.143595>
- Tosepu, R., Gunawan, J., Effendy, D. S., Lestari, H., Bahar, H., & Asfian, P. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Science of the Total Environment*, 725, 138436. <https://doi.org/10.1016/j.scitotenv.2020.138436>
- Wang, M., Jiang, A., Gong, L., Lu, L., Guo, W., Li, C., & Li, H. (2020). Temperature significantly change COVID-19 transmission in 429 cities. *medrxiv*.
- Wang, Q., Dong, W., Yang, K., Ren, Z., Huang, D., Zhang, P., & Wang, J. (2021). Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors. *International Journal of Infectious Diseases*, 105, 675–685. <https://doi.org/10.1016/j.ijid.2021.03.014>
- Wong, D. W., & Li, Y. (2020). Spreading of COVID-19: Density matters. *PLoS ONE*, 15(12), e0242398. <https://doi.org/10.1371/journal.pone.0242398>
- World Tourism Organization. (2019). *International tourism highlights* (2019th ed.). UNWTO. <https://doi.org/10.18111/9789284421152>
- Xie, J., & Zhu, Y. (2020). Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of the Total Environment*, 724, 138201. <https://doi.org/10.1016/j.scitotenv.2020.138201>
- Yang, T. C., Kim, S., Zhao, Y., & Choi, S. W. E. (2021). Examining spatial inequality in COVID-19 positivity rates across New York City ZIP codes. *Health & Place*, 69, 102574. <https://doi.org/10.1016/j.healthplace.2021.102574>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Moreno, R., & Vayá, E. (2024). Geographical distribution of the COVID-19 pandemic and key determinants: Evolution across waves in Spain. *Geographical Research*, 62(4), 486–502. <https://doi.org/10.1111/1745-5871.12669>