RESEARCH ARTICLE

Health Economics WILEY

How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces

Luis Orea[1,2] | Inmaculada C. Álvarez[2,3]

[1]Department of Economics, School of Business and Economics, University of Oviedo, Oviedo, Spain

[2]Oviedo Efficiency Group, University of Oviedo, Oviedo, Spain

[3]Departamento de Análisis Económico: Teoría Económica e Historia Económica, Universidad Autónoma de Madrid, Madrid, Spain

Correspondence
Luis Orea, Department of Economics, School of Business and Economics, University of Oviedo, Avda. del Cristo s/n, 33006 Oviedo, Spain.
Email: lorea@uniovi.es

Abstract

This paper examines the propagation of COVID-19 across the Spanish provinces and assesses the effectiveness of the Spanish lockdown of the population implemented on March 14, 2020 in order to battle this pandemic. To achieve these objectives, a standard spatial econometric model used in economics is adapted to resemble the popular reproduction models employed in the epidemiological literature. In addition, we introduce a counterfactual exercise that allows us to examine the Gross domestic product (GDP) gains of bringing forward the date of the Spanish Lockdown. We find that the number of COVID-19 cases would have been reduced by 70.4% in the absence of spatial propagation between the Spanish provinces. We also determine that the lockdown prevented the propagation of the virus within and between provinces. As such, the Spanish lockdown reduced the number of potential COVID-19 cases by 82.8%. However, the number of coronavirus cases would have been reduced by an additional 11.6% if the lockdown had been brought forward to March 7, 2020. Finally, an earlier lockdown would have saved approximately 26,900,000,000 euros.

KEYWORDS
COVID-19, Spanish lockdown, spatial propagation

JEL CLASSIFICATION
I1, H840, Q54, R12

1 | INTRODUCTION

An outbreak of a new coronavirus disease that causes potentially lethal respiratory tract infections in humans was detected for the first time in China in December 2019. The so-called coronavirus (COVID-19) spread rapidly to other countries. The first wave of contagion of COVID-19, hit Europe hard, especially Italy, Spain and several weeks later the United Kingdom. In Spain, the first case was confirmed in the Canary Islands on January 31. As shown in Figure 1, the virus spread rapidly to other provinces due to human mobility. All the Spanish provinces had already registered cases by March 14. Social distancing was encouraged on 9 March. The Governments of Madrid, La Rioja and the Basque Country
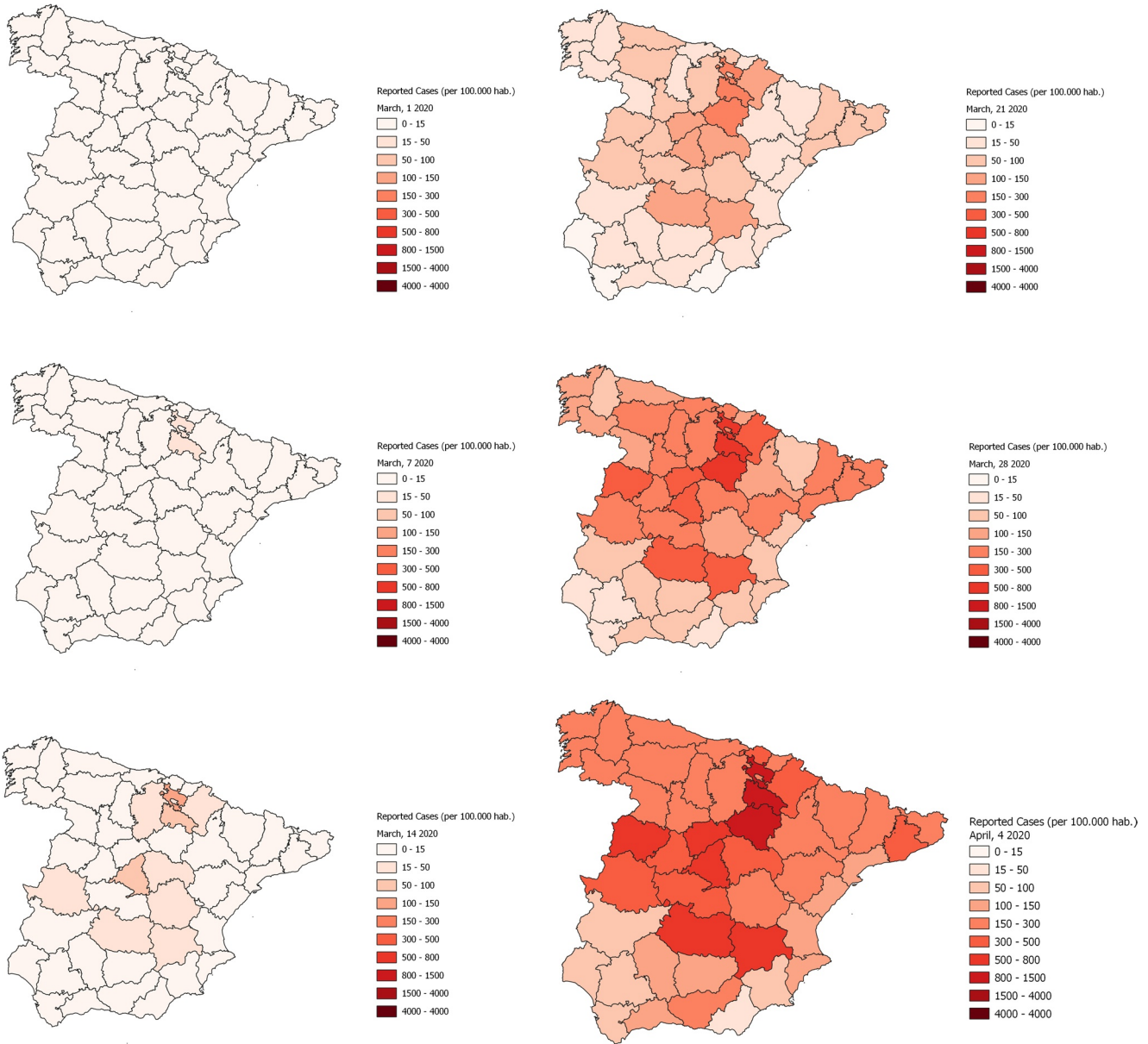
FIGURE 1 Geographical distribution of cumulative cases from March 1, 2020 to April 4, 2020 [Colour figure can be viewed at wileyonlinelibrary.com]

prohibited all in-class teaching in their regions over the following 3 days. Local outbreaks forced the Government of Cataluña to quarantine four Catalan municipalities on 12 March. The Spanish government declared a national lockdown of the population (or state of alarm) and prohibited public events on 14 March in an attempt to combat COVID-19. All shops except pharmacies and stores selling basic necessities were also forced to close. As the pandemic continued to spread after this date, it is germane to assess the effectiveness of this dramatic public intervention as well as the impact of other control measures. How human mobility explains the initial spread of COVID-19 is also an interesting issue worthy of close examination as it might prove helpful in understanding the propagation of the pandemic as well as limiting the impact of future waves.[1]

This paper aims to shed some light on the above issues using a spatial econometric analysis of the coronavirus propagation in Spain. Our empirical model aims to explain the daily evolution of the confirmed cases in the Spanish mainland provinces during the period between the onset of the pandemic in each province and April 4, 2020. In line with Giuliani et al. (2020), we distinguish between the propagation of the virus within a neighborhood, city or province and the propagation of COVID-19 across provinces. The origin of said spatial dimension of propagation is the high mobility of people

across provinces. This feature enables us to test whether the lockdown was effective in both preventing the propagation of the coronavirus between provinces and in attenuating the propagation of the virus within each province.

The added value of this study is the following. This is the first paper that examines the effectiveness of the control measures in Spain, as well as being one of the first in the recent literature that achieves this objective by controlling for spatial propagation effects, an issue that is treated only marginally in the epidemiological literature. Noteworthy exceptions are Giuliani et al. (2020), Gross et al. (2020), and Dickson et al. (2020). While most of the previous literature aims to estimate reproductive numbers, mortality, and other epidemic features, we apply more standard econometric techniques used in economics to carry out our empirical exercise. We show how such a model can be adapted to resemble the popular reproduction-based models used in the epidemiological literature, which often ignore the existence of spatial propagation effects and unobserved local conditions.

In addition, this paper includes a second major contribution, given that we also examine the economic impact of the Spanish lockdown implemented on March 14, 2020 in terms of Gross domestic product (GDP) losses at regional level. Based on the annual GDP growth rate forecasts per week of lockdown provided by BBVA Research (2020), using a counterfactual exercise we compute the economic effect of the actual lockdown and the GDP gains of bringing forward the date of the Spanish Lockdown.

The paper is structured as follows. Section 2 summarizes the empirical strategy used in this paper to assess the effectiveness of sizable public control measures implemented nationwide in Spain aimed at containing the outbreak, controlling for (and measuring) expected propagation effects across the Spanish mainland provinces. Section 3 briefly describes the data used in the empirical analysis and its sources. Section 4 provides the parameter estimates and discusses the main results. Finally, Section 5 presents the conclusions.

2 | MODELING LOCKDOWN IMPACT AND CORONAVIRUS PROPAGATION

This section outlines the main features of the empirical strategy used in this paper to assess both the propagation of coronavirus across the Spanish provinces and the effectiveness of the control measures implemented at containing the outbreak. We also discuss in two separate sub-sections the drawbacks of our empirical strategy, as well as the choice of the most suitable econometric specification for achieving the abovementioned objectives.

2.1 | Epidemic curve specification

This sub-section introduces the functional form of the epidemic curve to be estimated and the set of variables that will be used to capture the spread of the virus within and between provinces.

Consider a panel of $i = 1,\ldots,n$ provinces observed on $t = 1,\ldots,T$ days. Let E_i denote the *onset date* of the epidemic, that is, the date in which province i reports its first coronavirus case. We then analyze the development of the epidemic in each province, that is, the temporal evolution of coronavirus cases once each province reports its first coronavirus case.

Let Y_{it} denote the cumulative number of confirmed (reported) coronavirus cases until day t in province i. As is customary in panel data settings, we next assume that the number of cases in day can be expressed as a function of the number of cases on a previous day as follows:

$$Y_{it} = \beta_{it} Y_{it-1} \tag{1}$$

where β_{it} can be interpreted as a heteroskedastic autoregressive parameter. For ease of notation, we have chosen a single temporal lag of Y_{it} to represent this relationship. The autoregressive model (1) can be viewed as a *reduced-form* model that simply aims to fit the observed epidemic curve of cumulative cases. Therefore, our model does not make assumptions on the underlying parameters that determine the contagion of COVID-19. In this sense, we later show that a linear specification of Equation (1) can fit the epidemic curve of several epidemiological models, which differ in their assumptions on the incubation period, and other critical parameters. For instance, the popular Susceptible, Infected and Recovered (SIR) and Susceptible, Exposed, Infectious and Recovered (SEIR) epidemiological models yield time-varying growth rates of cumulative cases, regardless of whether daily or longer temporal lags are used.

In this sense, a key variable to carry out this analysis is the *epidemic time* $K_{it} = t - E_i$, which denotes the number of days relative to the onset date. We expect that the rate of growth of coronavirus cases varies with K_{it} as the traditional epidemic curve for a single wave has an S-shaped form.

The key aim of the coronavirus control measures is to reduce β_{it}. If β_{it} is equal to one, there are no new infections, and the pandemic has therefore been controlled. If β_{it} is greater than unity, new infections have been reported and the coronavirus pandemic is still spreading among the population despite the efforts to prevent the propagation of the virus. Our beta parameter (i.e., the rate of growth of cumulative cases) thus plays the same role as the so-called "reproductive number of the infection" (R), a fundamental epidemiological quantity, representing the average number of infections per infected case over the course of their infection. As we will show later, our beta parameter is also related to another commonly used epidemiological quantity: the so-called "growth rate," which is often defined as the proportional (per capita) change in number of new cases per unit of time.

In order to obtain a simple empirical specification of Equation (1), we take natural logarithms and perform a first differentiation of the model. This yields the following expression:

$$lnY_{it} - lnY_{it-1} = ln\beta_{it} = e^{\alpha_i + \gamma Z_{it} + \lambda W_i X_t} \tag{2}$$

where α_i is a set of province-specific but time-invariant fixed effects, $ln\beta_{it}$ is an exponential function of a set of covariates in order to impose the theoretical restriction $\beta_{it} \geq 1$, $X_t = \left(X_{1t}, X_{2t}, \ldots, X_{Nt}\right)$ is a $Nx1$ vector of explanatory variables of the Spanish provinces, and $W_i = \left(W_{i1}, W_{i2}, \ldots, W_{iN}\right)$ is a spatial weight vector where the weights ($W_{in} > 0, \forall i \neq n$) measures the degree of human mobility (connectivity) between provinces. The λ parameter is the spatial autoregressive coefficient that measures the degree of spatial correlation between provinces. In our application, it can be interpreted as the propagation effect caused by the mobility of people across provinces.

The vector of covariates Z_{it} includes two sets of variables. Z_{it} firstly includes a third-order function of lnK_{it} in order to capture the temporal pattern of the virus epidemic, conditional on the set of control measures.[2] The growth rate of cumulative cases in a simple SIR epidemiological model changes (decreases) with K_{it} either in levels or logs (see the temporal evolution of the growth rate of cumulative cases provided in Figure A1 of Appendix A that can be obtained replicating the same simulation of Chudik et al., 2020). The decline in growth rates for this model is not linear over time, which in turn explains why the epidemic curve of cumulative cases is S-shaped. Similar comments can be made if we introduce an incubation period into the SIR model obtaining an SEIR model (see, e.g., Institute for Disease Modelling, 2020). Moreover, if the model is deterministic, the simulated growth rates of cumulative cases can be predicted accurately using a third-order function of lnK_{it} (see again Figure A1 in Appendix A).

As pointed out by a referee, the time-varying growth rate of cumulative cases decreases in an SIR model probably because, in this model, a higher percentage of the population is no longer susceptible to the virus as time passes. Obviously, in a more realistic model, the growth rate of cumulative cases might also change over time due to the importation of cases (from other provinces or geographical areas) and the introduction of non-pharmaceutical interventions. On the one hand, these phenomena might explain why the S-shape of the epidemic curve cannot be perceived visually and on the other hand, justifies the inclusion of other explanatory variables such as spatially lagged indicators of the pandemic in neighboring provinces and dummy variables to capture the Spanish lockdown.

Second, Z_{it} includes a dummy variable $M14_t$ that takes the value 1 from March 14, 2020, the day marking the imposition of most of the coronavirus control measures by the Spanish Government. We also include 1- and 2-week lags of this dummy variable (i.e., $M21_t$ and $M28_t$) in order to capture larger effects attributable to the lockdown as time passes. This is an expected result due to the gap, which exists between when a person becomes infected and when they might subsequently infect another person, which is on average about 6 or 7 days (see, Flaxman et al., 2020, p. 18). Moreover, as pointed out by a referee, this result might also be caused by the lag between infection and the onset of symptoms and the existence of a large proportion of under-reported cases due to testing in March being saved and prioritized for only the most severe hospital cases.

Notice that our model specification looks like a Difference-in-Difference (DiD) model where we compare an outcome variable before and after treatment (a policy measure), having controlled for unobserved differences across units (provinces). Although the lockdown of the population in Spain was implemented in all provinces on March 14, 2020, the advance of the pandemic in each province was rather different at that time. Therefore, our identification strategy is based on the relatively large dispersion of pandemic developments (i.e., onset dates) across provinces, and that the onset dates are orthogonal to the lockdown implementation date.

We estimate the above model *after* taking natural logarithms to make it linear. Once we take natural logarithms, and a traditional noise term is added, the model to be estimated is:

$$lnRate = \alpha_i + \gamma Z_{it} + \lambda W_i X_t + v_{it} \tag{3}$$

where Rate $= lnY_{it} - lnY_{it-1}$, v_{it} is a mean-zero error term capturing random shocks, measurement or specification errors, and other unobservable variables not correlated with the rates of growth determinants. We used the logarithm transformation of the growth rates because it can be estimated using the standard linear Fixed-Effect (FE) estimator, which is equivalent to a linear panel data DiD estimation (Lechner, 2010, p. 189). This estimator ensures obtaining consistent causal effects attributable to a given policy measure, even in those cases where the time-invariant unobservable variables are correlated with the treatment variable (the lockdown dummy variable in our case). For instance, in our application, we might think that the centrality of Madrid and the greater mobility of the people living in Madrid and other populated cities/provinces were responsible for triggering the implementation of the Spanish lockdown.

It is also worth mentioning that in our paper we are not examining causal *epidemiological* effects in the sense that, for instance, infected individuals in period t cause secondary infections in period $t + 1$, and so on. This type of causal effect cannot be examined using a reduced-form model that simply aims to fit the observed epidemic curve of cumulative cases. However, the DiD specification of our reduced-form model is able to measure causal effects of a different nature, that is, those attributable to the public control measures implemented nationwide in Spain around March 14, 2020 aimed at containing the coronavirus outbreak during the first wave of the pandemic.

There is an extensive literature on human mobility for measuring the spread of infectious diseases. In this sense, it is worth mentioning the articles by Belik et al. (2011) and Bajardi et al. (2011), among others, that provide computational and theoretical models seeking to address the effect of human mobility and mobility restrictions on containing outbreaks of infectious diseases. Findlater and Bogoch (2018) find that the increasing volume of passenger travel, especially by air, enabled the global epidemic transmission. More recently, the use of new technologies such as mobile phones has facilitated the measurement of human mobility and its effects on disease connectivity (Lai et al., 2019). The researchers actually focus on severe acute respiratory syndrome coronavirus 2, concluding that human mobility predicts the spread and size of the epidemic and that travel restrictions are particularly useful in the early stage of the outbreak (see, e.g., Kraemer et al., 2020). This literature also demonstrates that viruses can spread through human contact patterns (Liu et al., 2020), given that human mobility contributes to promote social interaction (Mollgaard et al., 2017). Several studies corroborate these findings for Europe (see, e.g., Iacus et al., 2020; Lemey et al., 2021).

Please note that we use a Spatial Lag Model (SLX) specification to examine the role of human mobility in spreading the virus across the Spanish provinces. Inter-provincial mobility is captured using the spatial weight matrix $W = (W_1, \& , W_N)$. This spatial matrix can be computed in different ways. We follow Giuliani et al. (2020) and Gross et al. (2020) and use a contiguity or binary W matrix, where the weights equal one for adjacent units and zero for non-bordering units. In their spatial analysis of the spread of COVID-19 in Italy, Bourdin et al. (2021) performed several tests to select the best spatial weight matrix and selected, like us, the first-order contiguity matrix.

We select the epidemic time of neighboring provinces (i.e., $X_{it} = ln\,K_{it}$) in order to capture the potential propagation effects between provinces for two reasons. First, this variable is exogenous by construction. In a Spatial Autoregressive model (SAR) specification, X_{it} is replaced with (a transformation of) the dependent variable, which is endogenous and should thus be instrumented as long as good instruments are available. Second, Vega and Elhorst (2015, p. 342) suggest taking the SLX model as a point of departure because this is not only the simplest specification but is also more flexible in modeling spatial spillover effects than other specifications.

2.2 | Drawbacks

Three drawbacks of our empirical strategy are worth noting. First, although the linear FE model (3) has some features that are very appealing for our application, estimating the above logged linear model implies dealing with the zero growth rates of cumulative cases that often appear at the beginning of outbreaks. We can address this issue by dropping such observations from the sample. As this approach might generate some kind of sample selection bias if the missing observations are not random, we instead replace the zero values with a tiny but positive number before taking logs and keep the adjusted zero-value observations in our sample. We include a new dummy variable controlling for (adjusted) zero values as an additional explanatory variable. This variable not only allows us to control for potential measurement issues but also to prevent the observed sharp declines in growth rates caused by zero values to distort the third-order parametric function of epidemic times.

Second, in the first wave of the pandemic, no European country had sufficient testing capacity so that reported cases are a small fraction of the true number of infections. We can discuss whether this issue matters in our empirical application using the preliminary results of Orea et al. (2021), an ongoing study that complements the current paper as it tries to

account for the prevalence of undocumented cases. In this paper we propose a stochastic frontier analysis approach for estimating epidemic curves, where the unobserved cases are proxied using a one-sided random term in the same fashion as firms' inefficiency in production economics. We find that the average reporting rate is around 42%. Despite this, we obtain very similar effects due to lockdown on the growth rates of coronavirus cases (6.8 percentage points [pp] on average) compared to our non-frontier application. So, our results would seem to be quite robust in terms of this issue.

Another but related matter has to do with the onset date of the pandemic used in our paper. Our epidemic time variable is defined as the number of days relative to the *observed* onset date of the pandemic, which relies on reported cases. Therefore, it is not a necessary circumstance that a single reported case on a certain date seeded the pandemic in a particular province due to underreporting of cases. In order to see whether in practice the gap between observed and true onset dates is an important issue, we have modified the simulation of Chudik et al. (2020) and simulated several scenarios with different observed onset dates due to underreporting.[3] Two results of the simulation are worth mentioning. First, the goodness-of-fit of our model does not deteriorate when underreporting increases if the level of underreporting is common to all provinces. Second, the goodness-of-fit of the model does deteriorate when underreporting is large and the gap between observed and true onset dates varies notably across provinces. In this case, however, a linear model with fixed effects allowed us to retrieve the predictive capabilities of the model.

2.3 | Discussion on modeling choice

2.3.1 | Local versus global spatial spillovers

In this sub-section we discuss the nature of the spillovers generated by the SLX spatial specification of our epidemic curve. The spillovers induced by an SLX model are *local* in the sense that once the virus is transmitted from a province to another neighboring province, the transmission does not feedback and does not reverberate to other provinces. In this case, only adjacent neighbors are involved, but not higher-order neighbors. In contrast, the SAR model yields a more global spillover effect because it assumes that an impact on neighboring provinces reverberates to the neighbors of the neighboring provinces, neighbors to the neighbors, and so on, thus generating endogenous interaction and feedback effects (see LeSage, 2014). In this case, the propagation of an original outbreak involves more spatial observations.

The epidemiology literature focusing on the spatial propagation of COVID-19 highlights the contribution to the spread of the virus of both cross-border travel (Lemey et al., 2021) and local transmission (du Plessis et al., 2021). However, these papers do not discuss explicitly whether their transmission channels do have feedback effects between geographical units. This is the key issue that should guide the selection of a spatial econometric model. Although we believe that most of the inter-provincial mobility is local in nature due to regular commuting, we cannot rule out the possibility of more global effects caused by the transportation of goods or by business and leisure travelers.

As we do not have a theoretical justification for the selected spatial specification, we will proceed as follows with our empirical application. First, we will verify that the SLX model is able to capture all the spatial dependence in the dependent variable through a set of spatial autocorrelation tests on the model's residuals. We will next provide the parameter estimates of an SLX model that uses a *W* matrix defined using information on human mobility across *all* Spanish provinces, that is, not only between adjacent provinces. In this case, more spatial observations are involved, as occurs in the SAR and spatial Durbin models.

2.3.2 | Linear versus count regression models

In this sub-section we discuss the advantages of using a linear model instead of a count model. Both models mainly differ in their dependent (outcome) variables and distributional assumptions.[4] Despite these differences, the parameter estimates in our linear model can be interpreted as a semi-elasticity of the number of new cases with respect to an explanatory variable, in the same fashion as in count regression models.[5]

Although the interpretation of the estimated parameters is the same, the linear specification has some features that are critical in our application in order to measure the effectiveness of the Spanish lockdown in containing the propagation of COVID-19. First, running a linear model allows us to estimate a DiD model using the traditional FEs estimator. Estimating a DiD model using a count regression model is contentious as different empirical strategies exist for incorporating fixed effects into a count regression model, and some of them are not true FEs models (see Allison & Waterman, 2002).

Moreover, Lechner (2010, p. 196) shows that estimating a DiD model with the standard specification of a count regression models (and other popular nonlinear models) would usually lead to an inconsistent estimator. Second, as the growth rate of cumulative cases is much less volatile than the number of new cases (or its growth rate), our linear model provides more accurate predictions than a count model. This is a feature of the model that is important in our application because we use predicted values to carry out our counterfactual analyses aimed at examining the effect of the Spanish lockdown.

Despite the fact that the FE linear model has some features that are very appealing for our application, we also provide the parameter estimates of a Negative Binomial (NB) model for robustness analyses. The NB model is also estimated using two different W matrices, in the same fashion as the linear models. Whereas the contiguity-based W matrix is computed using binary values indicating adjacent provinces, the so-called mobility-based W matrix is computed using information on human mobility across all the Spanish provinces.

3 | SAMPLE AND DATA

We have used several sources in order to collect a province-based dataset of coronavirus cases that permits the use of spatial econometric techniques in order to capture spatial propagation effects across Spain. As most control measures began on the days of March 13, 2020 and March 14, 2020, we analyze data on coronavirus cases 2 weeks before and 2 weeks after those dates. In particular, our dataset covers the period between the onset of the pandemic in each province and April 4, 2020.

The daily evolution of laboratory-confirmed COVID-19 cases in Spanish mainland provinces was collected manually by the authors from the official press releases of the Spanish regional governments, the Ministry of Health and Wikipedia. In particular, we had to consult these information sources to extend backward the provincial data published by *Datadista* in GitHub under a free License since March 13, 2020,[6] the latter source extracting their data from a variety of documents published by the Ministry of Health. From March 28, 2020 onward, we collected the data directly using *RTVE Flourish*.[7] We used the regional online data released by the Ministry of Health[8] and the province-level data released by the Spanish regional governments in order to correct typos and the lack of information on coronavirus cases in some provinces (e.g., in Galicia). It should be noted that we were unable to obtain province-level data for the Cataluña region. For this reason, the whole region is treated as a single province.

We do not show the temporal evolution of reported coronavirus cases in each province due to space limitations, but they can be found in Orea and Álvarez (2020). We instead show the onset pandemic dates for each province in Figure 2, the latter determining the values of the epidemic times. A feature worth highlighting is the relatively large dispersion of onset dates across provinces. This feature is crucial for the estimation of Equation (3) because we need observations with
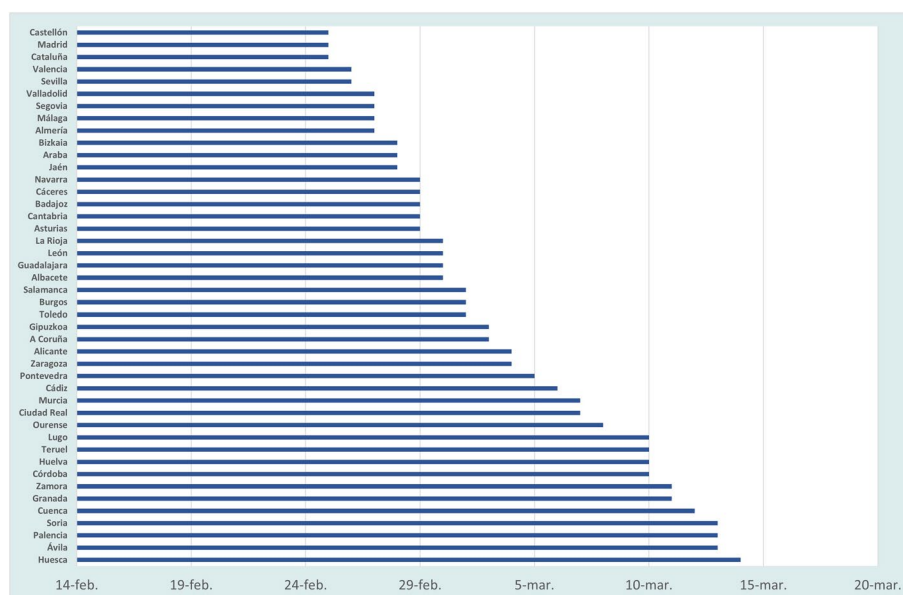

FIGURE 2 Observed onset date [Colour figure can be viewed at wileyonlinelibrary.com]
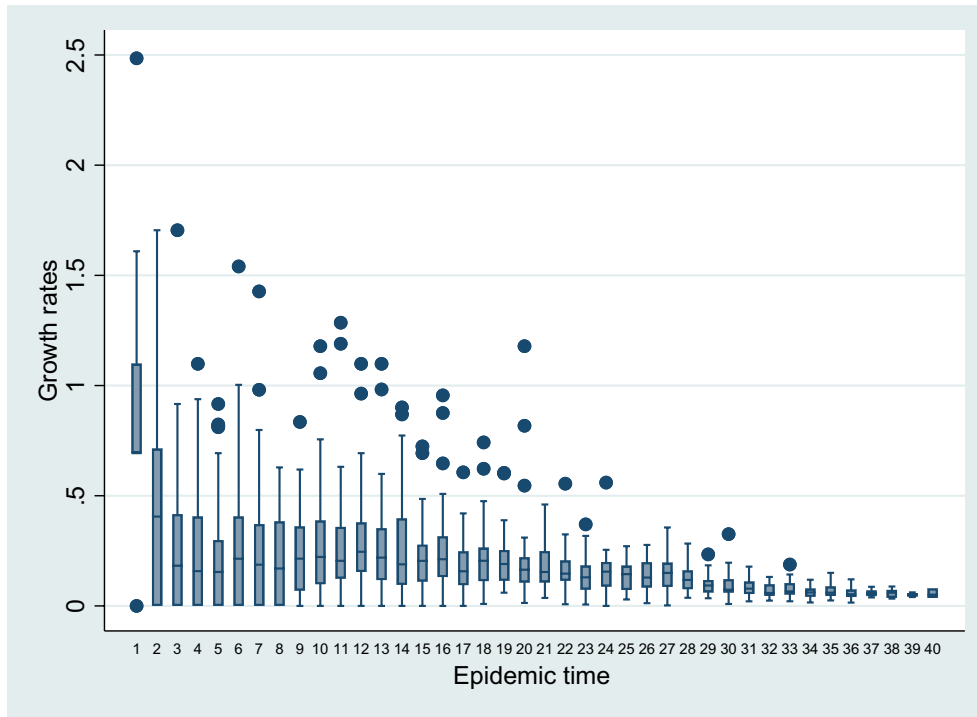
—WILEY—

FIGURE 3 Daily growth rates of cumulative cases [Colour figure can be viewed at wileyonlinelibrary.com]

both small and large epidemics in order to appropriately estimate the parametric function of lnK_{it}, especially before the lockdown implementation date.

Figure 3 shows the box-plots of the growth rates of cumulative cases by epidemic time. This figure clearly reveals two relevant features. First, the growth rates are much larger at the beginning of the pandemic than when the epidemic had progressed. That is, our dependent variable tends to decrease over the *epidemic* time. Second, the volatility is much larger when K_{it} is small and much smaller when K_{it} increases. This calls for using heteroskedasticity robust standard errors when estimating our models.

Both linear and NB models are also estimated using a spatial W matrix that is computed using information on human mobility across all the Spanish provinces. Data on mobility flows is obtained from the Spanish National Statistics Institute (INE), which in November 2019 initiated an ambitious project aimed at measuring daily mobility based on tracking spatial-temporal mobile position data.[9]

4 | EMPIRICAL RESULTS

4.1 | Parameter estimates

Table 1 shows the parameter estimates of several epidemic curves. Whereas the dependent variable using a linear model is the growth rate of cumulative cases, the new cases per day is the dependent variable using the NB model. All models have been estimated using the FE estimator because we reject that no correlation exists between the province-specific effects and the regressors using the traditional Hausman test at any significance level. All specifications in Table 1 provide very similar results, indicating that our empirical strategy is quite robust. The coefficients of the third-order function of lnK_{it} are all statistically significant. This is an expected result as the traditional epidemic curve is S-shaped and this form requires estimating up to a third-order function of the epidemic time.

The coefficients of $M14$, $M21$ and $M28$ allow us to test whether the Spanish lockdown and the previous control measures enacted by regional governments were successful in attenuating the spread of the virus *within* each province. As social distancing was encouraged on 9 March and in the following 3 days several regional governments prohibited in-class teaching and forced local quarantines, we find a statistically significant coefficient for $M14$. We also find a statistically significant coefficient for $M28$, an expected result due the national lockdown of the population. Figure 4 depicts the

TABLE 1 Parameter estimates

Dependent variable W matrix	SLX model ln Rate				Negative binomial New cases			
	Contiguity		Mobility		Contiguity		Mobility	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Epidemic time function								
lnK	−1.183***	0.138	−1.189***	0.141	−0.806***	0.260	−0.867***	0.252
lnK^2	0.514***	0.094	0.516***	0.099	0.613***	0.145	0.620***	0.151
lnK^3	−0.110***	0.020	−0.106***	0.022	−0.039	0.034	−0.037	0.035
Lockdown variables								
$M14$	−0.390***	0.089	−0.308***	0.088	−0.414**	0.170	−0.145	0.120
$M21$	−0.039	0.064	−0.070	0.068	0.129	0.101	0.139	0.106
$M28$	−0.330***	0.064	−0.362***	0.068	−0.546***	0.100	−0.542***	0.105
SLX variables								
$WlnK$	0.289***	0.057	0.032***	0.009	1.048***	0.186	0.140***	0.017
$WlnK \cdot M14$	−0.204***	0.036	−0.022***	0.006	−0.683***	0.084	−0.090***	0.009
Overdispersion parameter								
$ln\alpha$	-	-	-	-	−1.058***	0.104	−1.070***	0.103
Fixed effects	Yes	-	Yes	-	Yes	-	Yes	-
Zero-value effect	Yes	-	Yes	-	No	-	No	-
Lockdown effect (p.p.)	6.4	-	5.5	-	9.9	-	3.9	-
R-squared (%)	81.0	-	80.9	-	20.3	-	20.5	-
Obs	1411	-	1411	-	1411	-	1411	-

Note: $M14$, March 14; $M21$, March 21; $M28$, March 28; p.p., percentage points; K, epidemic time (number of days relative to the onset date).

Abbreviation: SLX, Spatial Lag Model.

*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

new cases per day over time for the different provinces, sorted by regions. This explains why each plot in Figure 4 includes multiple lines. There are two vertical red lines. Whilst the left one identifies the implementation of the Spanish lockdown (i.e. March 14, 2020), the right vertical line labels March 28, 2020. Notice that the daily incidence peaked around March 28, 2020 (i.e., when $t = 37$ or so), in many of the Spanish provinces. Therefore, this figure seems to support the idea that the Spanish lockdown started to have a significant effect on new cases and, hence, on cumulative cases 2 weeks after the implementation of the Spanish lockdown.

All models in Table 1 include two spatially lagged epidemic time variables. Our SLX spatial specification seems to capture all the spatial dependence in the dependent variable as we cannot reject the null hypothesis that the SLX residuals are not spatially correlated.[10] A key result of our empirical exercise is the positive and statistically significant coefficient found for the spatially lagged variable, $W_i \ln K_t$. This result provides evidence supporting the belief that human mobility did spread the virus across the country as it indicates that the growth rates of COVID-19 cases in one province depend on the development of the pandemic in other provinces.

Please note that we have interacted $M14$ with $W_i \ln K_t$. This implies that the coefficient of $W_i \ln K_t$ measures propagation effects before the implementation of the Spanish lockdown. The coefficient of $W_i \ln K_t \cdot M14$ is negative and statistically significant, indicating that the lockdown has attenuated the COVID-19 propagation between provinces. Moreover, the combined effect of $W_i \ln K_t$ and $W_i \ln K_t \cdot M14$ is close to zero in most models. This suggests that the lockdown has been quite effective in preventing the propagation of the coronavirus between provinces. In addition to this, the negative effect found for $W_i \ln K_t \cdot M14$ indicates that the lockdown has been more effective in provinces that are either close to the epicenters of the coronavirus or adjacent to provinces at a more advanced stage of the pandemic. As in our paper, Dickson et al. (2020) find that in the northern Italian provinces the Government containment measures not only succeeded in drastically reducing the transmission of COVID-19 amongst individuals within the Italian provinces, but also avoided contagions between neighboring areas.
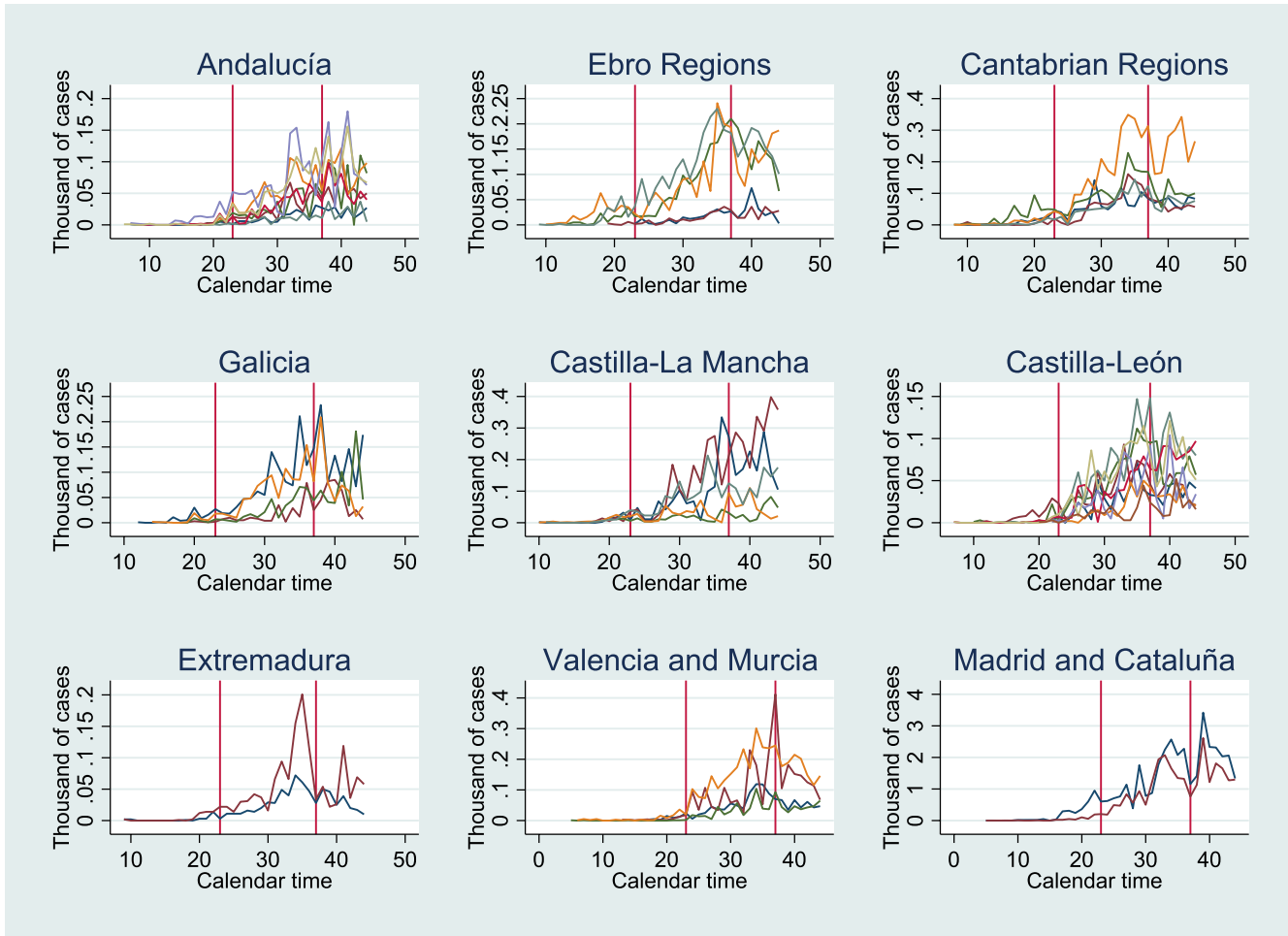
FIGURE 4   Provincial daily incidence, sorted by regions [Colour figure can be viewed at wileyonlinelibrary.com]

To conclude this section, it is germane to mention that we also regressed the estimated province fixed effects against a set of covariates in order to identify province-specific factors which intensify the pandemic's development in each province.[11] This information can be very useful for policy makers and health authorities when planning the relaxation of future lockdowns. We find that the most-populated provinces have suffered more acutely from COVID-19, probably due to the agglomeration of individuals and the more frequent use of public transport in these provinces. Coronavirus proved more intensive in those provinces with a relatively large share of highly educated workers. This result is most probably linked to provincial international connectivity and the probability of traveling abroad and/or importing cases of COVID-19 from other countries. We also found that the COVID-19 pandemic proved more severe in those provinces with a relatively large share of service sector workers. In contrast, the pandemic was less harsh in provinces with a relatively large share of workers in the agriculture and construction sectors. The risk of contagion in the service sector is not surprisingly much higher than in the construction and agricultural sectors because many service jobs are indoors, while most tasks in the other two sectors are mainly outdoors.

4.2 | Robustness analyses

Although in Table 1 all the models appear to provide similar parameter estimates, we next discuss some subtle but interesting differences. First, whereas the estimated parameters are very similar regardless of whether we use linear or NB models, the goodness-of-fit of the linear models is fairly large (over 80%) compared to the NB models (around 20%), an expected result due to the large volatility of daily new cases. Second, both linear and NB models are estimated using contiguity and mobility-based W matrices. While the first one is computed using binary values indicating adjacent provinces, the second one is computed using information on human mobility across all the Spanish provinces. The results of these

two competing spatial specifications are very similar due to 77% of the variation of the weights of the mobility-based W matrix being explained by the binary values of the weights of the contiguity W matrix. For instance, the mobility-based SLX model only attributes a slightly smaller effect to the Spanish lockdown than our preferred model. As the goodness-of-fit is slightly larger using the contiguity linkages, the latter model is used to carry out our simulation exercises. Interestingly enough, we find that the effect attributable to the Spanish lockdown in the NB models varies considerably when we change the W matrix. This seems to indicate that the linear models are, in our application, more robust to the definition of the W matrix.

4.3 | Spatial propagation

Our preferred model indicates that on average the growth rates of cumulative cases increases 5.1 pp, from 17.1% to 22.2%, due to the spatial propagation between provinces. The spatial spillover varies over time. For instance, while the growth rate of cumulative cases attributable to inter-provincial propagation is on average about 8.8 pp before March 14, 2020, it decreases up to 3.3 pp after the implementation of the Spanish lockdown. This result again suggests that the lockdown was effective in preventing the propagation of the coronavirus between provinces.

We also performed a counterfactual exercise using the parameter estimates of our preferred model in order to simulate what would have happened on April 4, 2020 in the case of no spatial propagation between provinces. Table 2 provides the results of this simulation exercise. This table shows remarkable reductions in cumulative cases in the absence of spatial spillovers between provinces. The number of reported cases in the mainland Spanish provinces on April 4, 2020 was 126,859. This number would have decreased to 37,557 if we drop the propagation between provinces. Therefore, the number of COVID-19 cases would have been reduced by 70.4% in the absence of spatial spillovers between the Spanish provinces.

TABLE 2 Spatial effects: Reported and simulated cases (April 4, 2020)

Region	Province	Reported A	Simulated B	Difference (%) C = (A − B)/A
Andalucía	Almería	346	242	−29.9
Andalucía	Cádiz	846	380	−55.0
Andalucía	Córdoba	974	225	−76.9
Andalucía	Granada	1477	516	−65.1
Andalucía	Huelva	279	149	−46.4
Andalucía	Jaén	914	395	−56.8
Andalucía	Málaga	1863	778	−58.2
Andalucía	Sevilla	1602	461	−71.2
Aragón	Huesca	396	244	−38.3
Aragón	Teruel	371	103	−72.2
Aragón	Zaragoza	2409	424	−82.4
Asturias	Asturias	1605	726	−54.8
Cantabria	Cantabria	1441	337	−76.6
CLM	Albacete	2653	622	−76.6
CLM	Ciudad real	3854	673	−82.5
CLM	Cuenca	497	181	−63.5
CLM	Guadalajara	858	265	−69.1
CLM	Toledo	2169	532	−75.5
CyL	Ávila	679	280	−58.8
CyL	Burgos	985	178	−81.9
CyL	León	1261	310	−75.4
CyL	Palencia	472	232	−50.8

TABLE 2 (Continued)

Region	Province	Reported A	Simulated B	Difference (%) C = (A − B)/A
CyL	Salamanca	1659	653	−60.6
CyL	Segovia	1148	259	−77.5
CyL	Soria	803	392	−51.1
CyL	Valladolid	1403	365	−74.0
CyL	Zamora	339	161	−52.4
Cataluña	Cataluña	26,032	8093	−68.9
Extremadura	Badajoz	672	188	−71.9
Extremadura	Cáceres	1375	570	−58.5
Galicia	A Coruña	2180	1373	−37.0
Galicia	Lugo	565	194	−65.6
Galicia	Ourense	921	410	−55.5
Galicia	Pontevedra	1519	821	−46.0
La Rioja	La Rioja	2592	589	−77.3
Madrid	Madrid	37,584	9266	−75.3
Murcia	Murcia	1235	507	−58.9
Navarra	Navarra	3073	976	−68.3
País Vasco	Álava	2639	381	−85.6
País Vasco	Vizcaya	4489	1248	−72.2
País Vasco	Guipúzcoa	1500	543	−63.8
Valencia	Alicante	2627	917	−65.1
Valencia	Castellón	852	402	−52.8
Valencia	Valencia	3701	992	−73.2
SPAIN	-	126,859	37,557	−70.4

Note: A, Number of reported cases; B, number of simulated cases.

Figure 5 allows an examination of the geographical distribution of the estimated spatial effects if we compare the actual distribution of cases on April 4, 2020 (top map) with the distributions of cases that would have been observed on April 4, 2020 in our hypothetical scenario (bottom map). This figure suggests that the spatial effect varies across provinces. Indeed, we found that while the reduction of cases in the event of no spatial propagation is much larger in provinces that are either close to the epicenters of the coronavirus or adjacent to provinces at a more advanced stage of the pandemic, it is smaller in the other provinces.

4.4 | Lockdown effects

We find using the parameter estimates of our preferred model that the growth rates of coronavirus cases decrease, on average, 6.4 pp (from 28.6% to 22.2%) due to the Spanish lockdown.[12] As aforementioned, the reduction in the growth rate of cumulative cases attributable to the lockdown in provinces that are close to (far from) the epicenters of COVID-19 or adjacent to provinces at more advanced stages of the pandemic, are much larger (smaller) than the abovementioned average value.

To provide information about the effectiveness of the lockdown by provinces, we have carried out two new counterfactual exercises that simulate what would have happened in two different hypothetical scenarios. We first simulate the number of coronavirus cases if the lockdown had not been implemented around March 14, 2020. The counterfactual values were simulated from March 14, 2020 onward by adding the difference between simulated and predicted growth rates to the observed growth rates cumulative cases. The counterfactual values are then used to compute reductions in the number of coronavirus cases for each province and not only for the whole country as in Flaxman et al. (2020). The second counterfactual exercise tries to examine what would have happened if the lockdown had been implemented on March 7, 2020. This information can be very useful for policy makers and health authorities in the event of new outbreaks
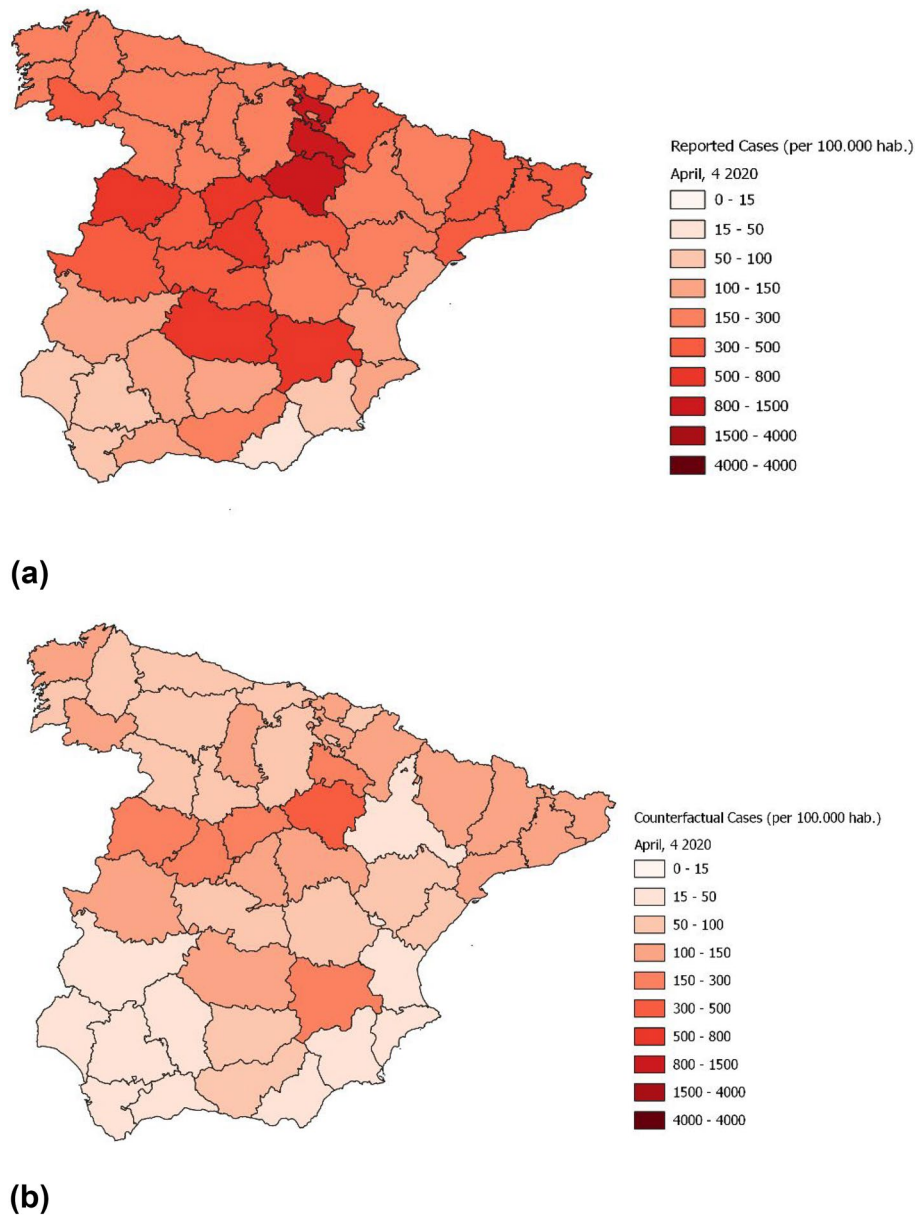
(a)

(b)

F I G U R E 5 Spatial spillovers: geographical distribution of cumulative cases on April 4, 2020. (a) Actual cases with spatial spillovers. (b) Counterfactual cases with no spatial spillovers [Colour figure can be viewed at wileyonlinelibrary.com]

of COVID-19 in Spain. The counterfactual values were simulated here from March 7, 2020 onward by subtracting the province-specific average of differences between simulated and predicted growth rates from the observed growth rates of confirmed cases. Table 3 provides the results of these two simulation exercises. Figure 6 compares the actual geographical distribution of coronavirus cases (shown in the middle map) with the counterfactual geographical distributions in the case of non-intervention (bottom map) and in the case of a hypothetical lockdown implemented on March 7, 2020 (top map).

The number of reported cases in Spanish mainland provinces on April 4, 2020 was 126,859. This number would have increased to 737,663 in the absence of lockdowns. Therefore, the lockdown implemented on March 14, 2020 reduced the number of potential COVID-19 cases by 82.8%. Similar numbers are found by Nussbaumer-Streit et al. (2020) in their rapid review of the literature related to COVID-19. They find that the quarantine measures reduce the number of people with the disease up to 81%. Using a similar approach, Cho (2020) recently found that the infection cases in Sweden would have been reduced by almost 75% had its policy makers followed stricter containment policies.

The largest reductions in coronavirus cases attributable to the Spanish lockdown are found again in provinces that are either close to the epicenters of the coronavirus or adjacent to provinces at more advanced stages of the pandemic, as

WILEY | 167

TABLE 3 Lockdown effects: reported and simulated cases (April 4, 2020)

Region	Province	Reported cases A	Simulated cases Mar 14 B	Simulated cases Mar 7 C	Difference (%) Mar 14 D = (B − A)/B	Difference (%) Mar 7 E = (A − C)/B	Difference (%) Both F = (B − C)/B
Andalucía	Almería	346	470	286	26.4	13.1	39.5
Andalucía	Cádiz	846	2021	380	58.1	23.1	81.2
Andalucía	Córdoba	974	15,078	35	93.5	6.2	99.8
Andalucía	Granada	1477	12,347	111	88.0	11.1	99.1
Andalucía	Huelva	279	671	103	58.4	26.3	84.8
Andalucía	Jaén	914	2617	404	65.1	19.5	84.6
Andalucía	Málaga	1863	4792	922	61.1	19.7	80.8
Andalucía	Sevilla	1602	9168	416	82.5	12.9	95.5
Aragón	Huesca	396	1073	90	63.1	28.6	91.7
Aragón	Teruel	371	4922	14	92.5	7.3	99.7
Aragón	Zaragoza	2409	37,372	169	93.6	6.0	99.6
Asturias	Asturias	1605	3431	902	53.2	20.5	73.7
Cantabria	Cantabria	1441	13,566	196	89.4	9.2	98.6
CLM	Albacete	2653	57,665	170	95.4	4.3	99.7
CLM	C. Real	3854	48,889	271	92.1	7.3	99.4
CLM	Cuenca	497	8106	11	93.9	6.0	99.9
CLM	Guadalajara	858	3913	241	78.1	15.8	93.9
CLM	Toledo	2169	24,228	266	91.0	7.9	98.9
CyL	Ávila	679	12,241	8	94.5	5.5	99.9
CyL	Burgos	985	10,829	117	90.9	8.0	98.9
CyL	León	1261	19,294	106	93.5	6.0	99.5
CyL	Palencia	472	2736	35	82.8	16.0	98.8
CyL	Salamanca	1659	5509	591	69.9	19.4	89.3
CyL	Segovia	1148	11,766	178	90.2	8.3	98.5
CyL	Soria	803	6145	39	86.9	12.5	99.4
CyL	Valladolid	1403	29,383	112	95.2	4.4	99.6
CyL	Zamora	339	1316	68	74.2	20.7	94.9
Cataluña	Cataluña	26,032	92,979	10,742	72.0	16.4	88.4
Extremadura	Badajoz	672	6332	97	89.4	9.1	98.5
Extremadura	Cáceres	1375	4685	534	70.6	18.0	88.6
Galicia	A Coruña	2180	3061	1698	28.8	15.8	44.6
Galicia	Lugo	565	3210	70	82.4	15.5	97.9
Galicia	Ourense	921	3210	235	71.3	21.4	92.7
Galicia	Pontevedra	1519	3050	820	50.2	22.9	73.1
La Rioja	La Rioja	2592	11,140	750	76.7	16.5	93.3
Madrid	Madrid	37,584	188,028	11,927	80.0	13.6	93.7
Murcia	Murcia	1235	2857	565	56.8	23.5	80.3
Navarra	Navarra	3073	18,269	747	83.2	12.7	95.9
País Vasco	Álava	2639	15,334	629	82.8	13.1	95.9
País Vasco	Vizcaya	4489	10,033	2433	55.3	20.5	75.8
País Vasco	Guipúzcoa	1500	3546	724	57.7	21.9	79.6

(Continues)

TABLE 3 (Continued)

Region	Province	Reported cases A	Simulated cases		Difference (%)		
			Mar 14 B	Mar 7 C	Mar 14 $D = (B - A)/B$	Mar 7 $E = (A - C)/B$	Both $F = (B - C)/B$
Valencia	Alicante	2627	5416	1376	51.5	23.1	74.6
Valencia	Castellón	852	2017	452	57.8	19.9	77.6
Valencia	Valencia	3701	14,950	1279	75.2	16.2	91.4
SPAIN	-	126,859	737,663	41,318	82.8	11.6	94.4
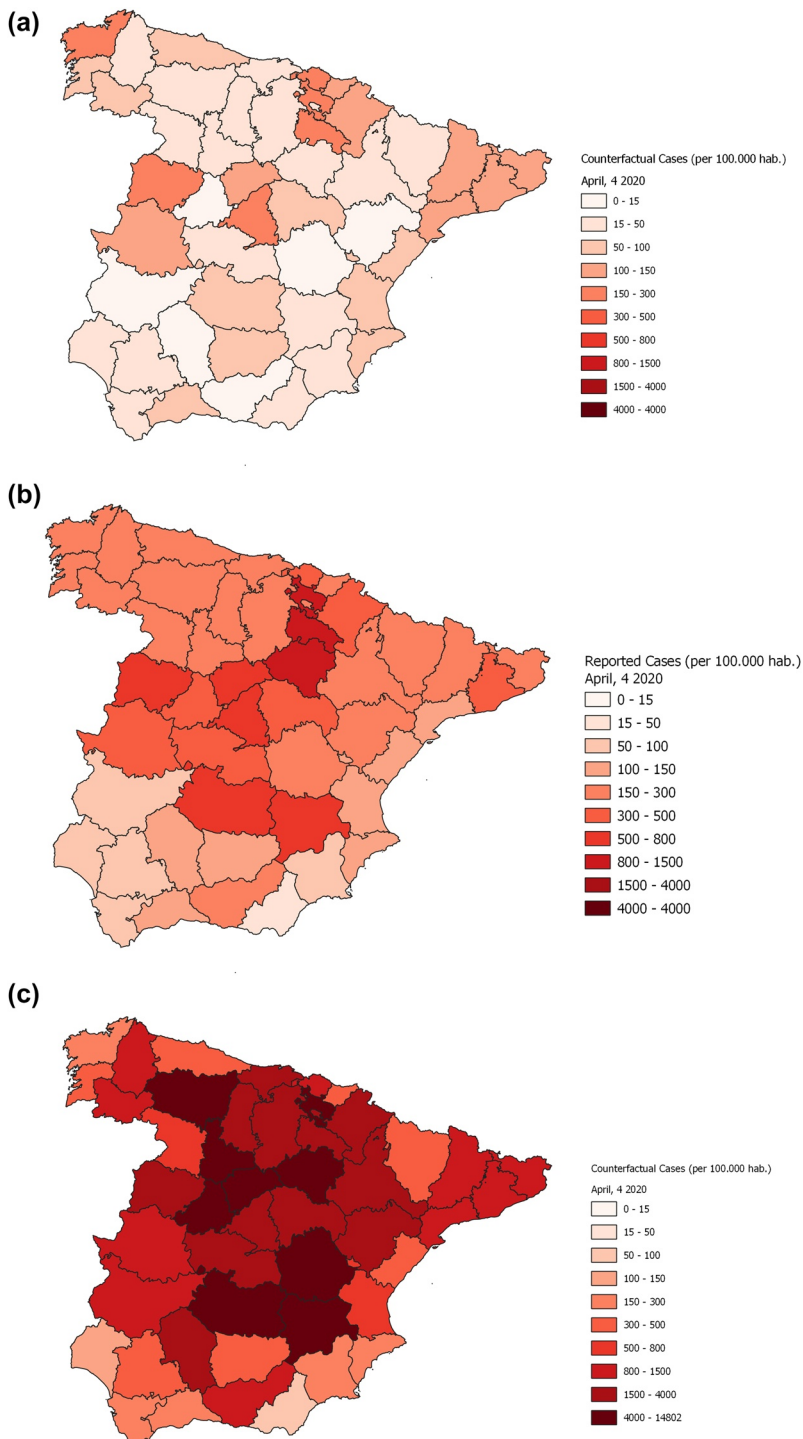
FIGURE 6 Lockdown effects: Geographical distribution of cumulative cases on April 4, 2020. (a) Counterfactual cases if the lockdown were implemented on March 7, 2020. (b) Actual cases with the lockdown implemented on March 14, 2020. (c) Counterfactual cases with no lockdown [Colour figure can be viewed at wileyonlinelibrary.com]

the two last maps in Figure 6 suggest. We next discuss what would have happened if the lockdown had begun on March 7, 2020. If the lockdown had been brought forward to March 7, 2020, the number of additional coronavirus cases would have been reduced from 126,859 to 41,318 in the Spanish Peninsula. Taken together both counterfactual analyses, the lockdown implemented on March 7, 2020 reduced the number of potential COVID-19 cases by 94.4%. Therefore, the number of coronavirus cases would have been reduced by an additional 11.6% if the lockdown had been brought forward to March 7, 2020, a reduction that potentially would have prevented the collapse of many hospitals in Spain.

4.5 | GDP savings of bringing forward the date of the Spanish lockdown

We finally examine the GDP gains of bringing forward the date of the Spanish lockdown. The second counterfactual exercise carried out in the previous section shows that many provinces would have had less than 28 new confirmed cases per 100.000 habitants during the 2 weeks prior to April 4, 2020. This in turn implies that on April 4, 2020 many regions would have already met one of the conditions stipulated by the Spanish Government to initiate the relaxation of the lockdown measures.

The easing of the lockdown restrictions in Spain began on May 11, except in Castilla-León, Cataluña and Madrid where it started on May 25. Table 4 shows the duration of the actual lockdown in the Spanish mainland regions (see the second column). The strictest part of the confinement lasted 8.3 weeks in all regions, except in the three aforementioned regions where the confinement was extended by 2 weeks. The first column in Table 4 shows the annual GDP growth rate forecasts per week of lockdown provided by BBVA Research (2020). Based on this information, the third and fourth columns use this information to compute the economic effect of the actual lockdown in terms of GDP growth rates and

TABLE 4   GDP losses of both simulated and actual lockdowns

	GDP growth per week[a] (in %)	Lockdown from 14 Mar onward			Simulated lockdown from 7 Mar onward			Difference in GDP losses[f] (in 1000 millions)
		Lockdown duration[b] (in weeks)	GDP growth[c] (in %)	GDP losses[d] (in 1000 millions)	Lockdown duration[e] (in weeks)	GDP growth[c] (in %)	GDP losses[d] (in 1000 millions)	
Andalucía	−1.04	8.3	−8.6	−13.4	5	−5.2	−8	5.3
Aragon	−0.85	8.3	−7.0	−2.5	4.5	−3.9	−1.4	1
Asturias	−1.00	8.3	−8.3	−1.9	6	−6	−1.4	0.5
Cantabria	−1.00	8.3	−8.3	−1.1	4	−4	−0.5	0.6
C. La Mancha	−0.80	8.3	−6.6	−2.6	6.2	−5.0	−2.0	0.6
Castilla Leon	−0.93	10.3	−9.5	−5.3	6.2	−5.8	−3.2	2.1
Cataluña	−1.03	10.3	−10.5	−23.3	8	−8.2	−18.2	5
Extremadura	−0.83	8.3	−6.8	−1.3	5.6	−4.6	−0.9	0.4
Galicia	−0.93	8.3	−7.7	−4.6	6.0	−5.5	−3.3	1.3
La Rioja	−1.00	8.3	−8.3	−0.7	8	−8	−0.7	0
Madrid	−0.93	10.3	−9.5	−21.1	8	−7.4	−16.4	5
Murcia	−1.04	8.3	−8.6	−2.6	4	−4.2	−1.3	1.4
Navarra	−0.91	8.3	−7.6	−1.5	8	−7.3	−1.4	0
País Vasco	−0.91	8.3	−7.6	−5.3	7.3	−6.7	−4.7	0.6
Valencia	−1.04	8.3	−8.6	−9.3	6	−6.2	−6.7	2.6
All regions	-	-	-	−96.6	-	-	−69.7	26.9

Abbreviation: GDP, Gross domestic product.

[a]Annual GDP growth rate per week of lockdown (BBVA Research forecast).

[b]The easing of the Spanish lockdown restrictions started on May 11, except in Castilla-León, Cataluña, and Madrid where it started on May 25.

[c]Annual GDP growth rate attributable to the whole lockdown.

[d]Losses computed using the GDP for 2019.

[e]Weighted average of provinces' lockdown duration.

[f]Simulated GDP losses minus GDP losses of the actual lockdown.

GDP losses. Given BBVA forecasts, the lockdown implemented on March 14, 2020 would have reduced Spanish GDP by 96,600 million euros.

The next four columns on Table 4 provide an estimate of the economic disruption corresponding to a hypothetical lockdown implemented on March 7, 2020. The lockdown of a province is assumed to start easing on April 11, 2020 if two conditions are satisfied. The first is that it meets the criterion mentioned above of having less than 28 new confirmed cases per 100,000 habitants during the 2 weeks prior to April 4, 2020. The second is a condition that has to do with the (relative) capacity of its health services to deal with new cases of COVID-19, which a province meets if it has less confirmed cases per capita than the median province on April 4, 2020. The lockdown would have lasted only 4 weeks if a province had met these two conditions on April 4, 2020. If only one condition is met, the easing of lockdown restrictions is assumed to start on April 25, 2020, in which case the lockdown would have lasted 6 weeks. Finally, if neither of the conditions is met, the easing of lockdown is assumed to start on May 9, that is, 2 weeks later. Once the duration of the lockdown has been simulated for each province, a weighted average is computed for the whole region using the relative GDP of each province as weights. The simulated regional lockdown durations are shown in the fifth column. The next two columns show the simulated annual GDP growth rate and GDP losses using the annual GDP growth rates per week of lockdown shown in the first column. Given the BBVA forecasts, our simulated lockdown implemented on March 7, 2020 would have reduced Spanish GDP by 69,700 million euros.

Finally, the last column on Table 4 shows the difference in GDP losses between the simulated and real lockdown. Summing across all regions, the estimated difference in GDP losses is around 26,900 million euros. Therefore, the simple economic analysis in Table 4 suggests that the final economic consequences of the confinement of population would have been much less severe if the Spanish lockdown had been brought forward to March 7, 2020.

5 | CONCLUSIONS

This paper examines the propagation of COVID-19 across the Spanish provinces and assesses the effectiveness of the Spanish lockdown of the population implemented on March 14, 2020 to combat the pandemic. To achieve these objectives, we use a spatial econometric model that somehow mimics the popular reproduction-based models used in the epidemiological literature.

The main findings of the paper are the following. We provide evidence supporting the belief that human mobility did spread the virus across the country given that we observe that the growth rate of COVID-19 cases in one province depends on the development of the pandemic in other provinces. We also find that the lockdown has been effective in both attenuating the propagation of the virus within each province as well as preventing the propagation of the coronavirus between provinces.

Our counterfactual analyses show that local and national lockdowns of the population are effective measures to combat COVID-19 in the absence of both pharmaceutical related measures (e.g., vaccines) and other non-pharmaceutical interventions (e.g., massive testing, face-masks available for the whole population, etc.). However, they should be implemented at the very early stages of the pandemic. On the one hand, our analyses suggest that carrying out a gradual relaxation of the control measures in Spain, both across provinces and sectors is preferable. On the other hand, we find that the GDP losses attributable to the confinement of the population would have been reduced by 26.9 thousand million euros if the Spanish lockdown had been brought forward to March 7, 2020. As such, we find that a rapid institutional response to the COVID-19 outbreak not only saves lives but would also have attenuated the economic impact of the Spanish coronavirus pandemic.

editing the English grammar of the text. We also gratefully acknowledge the data provided by Datadista in GitHub and Flourish in RTVE.

CONFLICT OF INTEREST
Luis Orea and Inmaculada C. Álvarez declare that they have no conflict of interest.

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

ORCID
Inmaculada C. Álvarez https://orcid.org/0000-0003-0495-623X

ENDNOTES
[1] We thank the reviewers for pointing out to us that human mobility is a proxy for what we really think effects the spread of a contagious disease, that is, human contact patterns.

[2] Separate analyses or more flexible models must be implemented in order to account for more than one contagion waves (see, e.g., Dickson et al., 2020, who use B-spline regressors to model complex nonlinear spatio-temporal dynamics in the propagation of the virus).

[3] We have modified the simulation of Chudik et al. (2020) as follows. We first generate the true evolution of total coronavirus cases in a representative province using the discrete-time SIR model developed by Chudik et al. (2020). Observed values for each province are then obtained by adjusting the theoretical values with simulated values for a one-sided (half-normal) random term capturing the proportion of undocumented cases. We replicate this procedure for different levels of underreporting. In all replications, the onset of the pandemic is associated with the day on which we observe the first case. Although all provinces have the same true onset date, the observed onset date of each province might differ due to underreporting.

[4] When the outcome is considered to be continuous, the data are frequently assumed to be normally distributed and linear least squares regression techniques are applied. The count regression models provide an alternative approach for the analysis of discrete data, provided that the outcome follows , for example, a Poisson distribution, the over dispersion issue is correctly specified, and the model adequately fits the data.

[5] Notice that $Rate_{it} \approx (Y_{it} - Y_{it-1})/Y_{it-1} = N_{it}/Y_{it-1}$, where N_{it} stands for new cases in day t in province i. As the conditional expectation of $lnRate_{it}$ in our model is $E[(lnN_{it} - lnY_{it-1}) \mid Z_{it}, W_i X_t] = \alpha_i + \gamma Z_{it} + \lambda W_i X_t$, the conditional expectation of lnN_{it} is $E(lnN_{it} \mid Z_{it}, W_i X_t, Y_{it-1}) = \alpha_i + \gamma Z_{it} + \lambda W_i X_t + lnY_{it-1}$. Therefore, despite their differences, both approaches are similar in the sense that the parameter estimates have the same interpretation.

[6] See https://github.com/datadista/datasets/tree/master/COVID%2019.

[7] See https://app.flourish.studio/visualisation/1451263/.

[8] See https://covid19.isciii.es/.

[9] See https://www.ine.es/experimental/movilidad/experimental_em.htm.

[10] According to Beenstock and Felsenstein (2019), the lack of spatial correlation in the residuals should be tested for each time period. The set of performed Moran's I and Geary's tests are available from the authors upon request.

[11] The parameter estimates are available from the authors upon request.

[12] Although the accumulated effect of this reduction is remarkable (see our results in Table 3), the epidemic did not stop growing by April 4, 2020. Using our model, we can only conclude that the Spanish lockdown helped to attenuate the COVID-19 propagation during the first wave of contagion.

REFERENCES
Allison, P. D., & Waterman, R. P. (2002). Fixed-effects negative binomial regression models. *Sociological Methodology*, *32*(1), 247–265. https://doi.org/10.1111/1467-9531.00117

Bajardi, P., Poletto, C., Ramasco, J. J., Tizzoni, M., Colizza, V., & Vespignani, A. (2011). Human mobility networks, travel restrictions and the global spread of 2009 H1N1 pandemic. *PLoS One*, *6*, e16591. https://doi.org/10.1371/journal.pone.0016591

BBVA Research. (2020). *Regional analysis Spain. Second quarter 2020*. https://www.bbvaresearch.com/en/publicaciones/regional-analysis-spain-second-quarter-2020/

Beenstock, M., & Felsenstein, D. (2019). *The econometric analysis of non-stationary spatial panel data*. Springer International Publishing.

Belik, V., Brockmann, D., & Geisel, T. (2011). Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review*, *X*(1), 011001. https://doi.org/10.1103/PhysRevX.1.011001

Bourdin, S., Jeanne, L., Nadou, F., & Noiret, G. (2021). Does lockdown work? A spatial analysis of the spread and concentration of Covid-19 in Italy. *Regional Studies*, 55, 1182–1193. https://doi.org/10.1080/00343404.2021.1887471

Cho, S. W. (2020). Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. *The Econometrics Journal*, 23(3), 323–344. https://doi.org/10.1093/ectj/utaa025

Chudik, A., Pesaran, M. H., & Rebucci, A. (2020). *Voluntary and mandatory social distancing: Evidence on COVID-19 exposure rates from Chinese provinces and selected countries*. NBER Working Paper 27039. http://www.nber.org/papers/w27039

Dickson, M. M., Espa, G., Giuliani, D., Santi, F., & Savadori, L. (2020). Assessing the effect of containment measures on the spatio-temporal dynamic of COVID-19 in Italy. *Nonlinear Dynamics*, 101(3), 1833–1846. https://doi.org/10.1007/s11071-020-05853-7

du Plessis, L., McCrone, J. T., Zarebski, A. E., Hill, V., Ruis, C., Gutierrez, B., Raghwani, J., Ashworth, J., Colquhoun, R., Connor, T. R., Faria, N. R., Jackson, B., Loman, N. J., O'Toole, Á., Nicholls, S. M., Parag, K. V., Scher, E., Vasylyeva, T. I., Volz, E. M., ... Pybus, O. G. (2021). Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science*, 371, 708–712. https://doi.org/10.1126/science.abf2946

Findlater, A., & Bogoch, I. I. (2018). Human mobility and the global spread of infectious diseases: A focus on air travel. *Trends in Parasitology*, 34(9), 772–783. https://doi.org/10.1016/j.pt.2018.07.004

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Perez-Guzman, P. N., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Boonyasiri, A., Boyd, O., Cattarino, L., ... Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584, 257–261. https://doi.org/10.1038/s41586-020-2405-7

Giuliani, D., Dickson, M. M., Espa, G., & Santi, F. (2020). Modelling and predicting the spatio-temporal spread of COVID-19 in Italy. *BMC Infectious Diseases*, 20(1), 700. https://doi.org/10.1186/s12879-020-05415-7

Gross, B., Zheng, Z., Liu, S., Chen, X., Sela, A., Li, J., Li, D., & Havlin, S. (2020) Spatio-temporal propagation of COVID-19 pandemics. *medRxiv preprint*. https://doi.org/10.1101/2020.03.23.20041517

Iacus, S. M., Santamaria, C., Sermi, F., Spyratos, S., Tarchi, D., & Vespe, M. (2020). Human mobility and COVID-19 initial dynamics. *Nonlinear Dynamics*, 101(3), 1901–1919. https://doi.org/10.1007/s11071-020-05854-6

Institute for Disease Modelling. (2020). *SEIR and SEIRS models*. https://docs.idmod.org/projects/emod-hiv/en/latest/model-seir.html

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493–497. https://doi.org/10.1126/science.abb4218

Lai, S., Farnham, A., Ruktanonchai, N. W., & Tatem, A. J. (2019). Measuring mobility, disease connectivity and individual risk: A review of using mobile phone data and mHealth for travel medicine. *Journal of Travel Medicine*, 26(3), taz019. https://doi.org/10.1093/jtm/taz019

Lechner, M. (2010). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3), 165–224. https://doi.org/10.1561/0800000014

Lemey, P., Ruktanonchai, N., Hong, S. L., Colizza, V., Poletto, C., den Broeck, F. V., Gill, M., Ji, X., Levasseur, A., Sadilek, A., Lai, S., Tatem, A., Baele, G., Suchard, M., & Dellicour, S. (2021). SARS-CoV-2 European resurgence foretold: Interplay of introductions and persistence by leveraging genomic and mobility data. This preprint is under consideration at a Nature Portfolio Journal. https://doi.org/10.21203/rs.3.rs-208849/v1

LeSage, J. (2014). What regional scientists need to know about spatial econometrics. *Review of Regional Studies*, 44(1), 13–32. https://doi.org/10.2139/ssrn.2420725

Liu, Y., Gu, Z., Xia, S., Shi, B., Zhou, X., Shi, Y., & Liu, J. (2020). What are the underlying transmission patterns of COVID-19 outbreak? An age-specific social contact characterization. *EClinicalMedicine*, 22, 100354. https://doi.org/10.1016/j.eclinm.2020.100354

Mollgaard, A., Lehmann, S., & Mathiesen, J. (2017). Correlations between human mobility and social interaction reveal general activity patterns. *PLoS One*, 12, e0188973. https://doi.org/10.1371/journal.pone.0188973

Nussbaumer-Streit, B., Mayr, V., Dobrescu, A. I., Chapman, A., Persad, E., Klerings, I., Wagner, G., Siebert, U., Christof, C., Zachariah, C., & Gartlehner, G. (2020). Quarantine alone or in combination with other public health measures to control COVID-19: A rapid review. *Cochrane Database of Systematic Reviews*, (4), CD013574. https://doi.org/10.1002/14651858.CD013574

Orea, L., & Álvarez, I. (2020). *How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces*. Working Paper 2020/3, FEDEA (Fundación de Estudios de Economía Aplicada). https://documentos.fedea.net/pubs/dt/2020/dt2020-03.pdf

Orea, L., Álvarez, I., & Wall, A. (2021). *Estimating the propagation of the COVID-19 virus with a stochastic frontier approximation of epidemiological models: A panel data econometric model with an application to Spain*. Efficiency Series Paper, 01/2021. Oviedo Efficiency Group, University of Oviedo. http://www.uniovi.es/oeg/ESP/esp_2021_01.pdf

Vega, S. H., & Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, 55(3), 339–363. https://doi.org/10.1111/jors.12188

APPENDIX A

FIGURE A1 Growth rates of cumulative cases in an Susceptible, Infected and Recovered models [Colour figure can be viewed at wileyonlinelibrary.com]