Extended abtract

EXTENDED ABSTRACT

Title: The impact of COVID-19 on mobility. Empirical evidence with a spatial seemingly unrelated regression model.

Authors and e-mails:

Fernando A López (Fernando.lopez@upct.es)
Román Mínguez (Roman.Minguez@uclm.es)
Jesús Mur (jmur@unizar.es)

Department: Métodos Cuantitativos, Ciencias Jurídicas y Lenguas Modernas

University: Universidad Politécnica de Cartagena

Subject area: *(Métodos para el análisis regional y urbano: big data, econometría espacial y sistemas de información geográfica)*

Abstract: *(minimum1500 words)*
There is broad consensus that this pandemic has been one of the most relevant events that has occurred during recent decades with impacts (social, medical, economic, etc.) that will last for a long time. However, in spite of the dramatic consequences, social distancing and the reduction of individual mobility are seen as the most effective strategies to slow down the spread of the disease. This position is shared by most experts and explains the prominent role assigned to both factors. Increasing mobility (or shortening social distance) leads to a greater incidence of the disease, and the contrary should be the result of reduced mobility (or greater social distance). This is the hypothesis that we want to explore. In particular, we focus on the case of Spain, using data from its provinces (NUTS3 in Eurostat terminology) during the first weeks of the pandemic. Spain, with 267,000 confirmed cases by the end of July 2020, is suffering intensely from the pandemic; these figures represented 16.34% of the cases in the EU28 and 1.83% of worldwide cases, while the equivalent percentages in terms of population were 9.40% and 0.63%, respectively.

In response to the rapidly growing number of COVID-19 cases, the Spanish Government (like many others) enacted orders to reduce individual mobility, declaring a State of Emergency (SE) on March 14. In practice, the SE implied a lockdown of the population in their homes. Many non-essential service activities, such as education, leisure and entertainment, commerce (except for food, health or electronics), etc. were suspended. On March 28, the halt also reached the sectors of manufacturing and construction, with grave implications for the Spanish economy. On April 10, both

sectors were allowed to resume their activities as there was a clear indication that the lockdown had been successful and, on June 23, the SE was lifted^[During the same time, the Spanish Government initiated a big data project whose purpose was to collect information about the mobility of the population by monitoring the location of all the cell phones in national territory (in fact, 80% of mobile phones are under control). The daily collection began on February 14 and includes all movements exceeding 500 meters. This information also distinguishes short range or intra-provincial trips whose origin and destination are both inside the same province, inter-provincial, or long range trips, where the province of origin differs from that of the destination.].

Despite the fact that mobility restrictions were dictated for the entire country, the Government's orders were unevenly followed in different provinces. Several factors can explain these regional differences. In the first place, provinces with high levels of essential services reduced mobility to a lesser extent than provinces with low levels of essential services. In the second place, as demonstrated by @Paez2020, something similar occurred in provinces with a high proportion of older people whose ability to move is more restricted. Lastly, in this period fear had an effect in the sense that the observed incidence of the disease dissuaded individuals from leaving their homes, thus reducing mobility. This perception changed from province to province because it depended on local conditions.
Our dataset consists of variables related to the two mobility indices (intra- and inter-provincial movements) for the 50 spatial units (provinces) and 17 temporal periods (weeks from February 14 to June 21) aimed at controlling the factors described above (see the help facility in the spsur package for details and formal definitions).

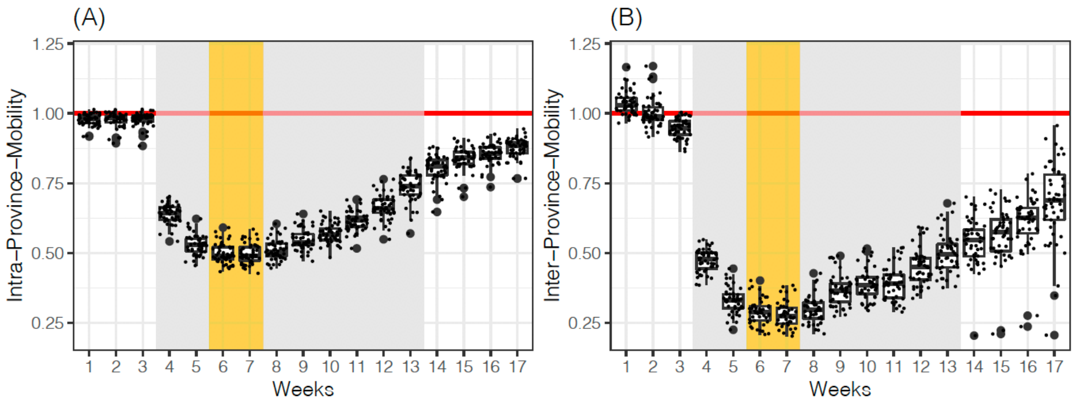

Figure 1 shows the weekly evolution of the two main variables involved in our research. Figures 1A and 1B refer to intra- and inter-provincial mobility. Note the strong reduction of the mobility indices, both intra and inter, produced in week 4 with respect to the pre-COVID reference week (February 14-21). This is not accidental since week 4 is the beginning of the State of Emergency (SE). Weeks 6 and 7 (in orange), correspond to the tightening of restrictive measures when only essential activities were allowed. After week 7, the confinement measures relaxed, almost reaching normality around week 13. In any case, the SE remained in effect until June 21, the end of week 17.

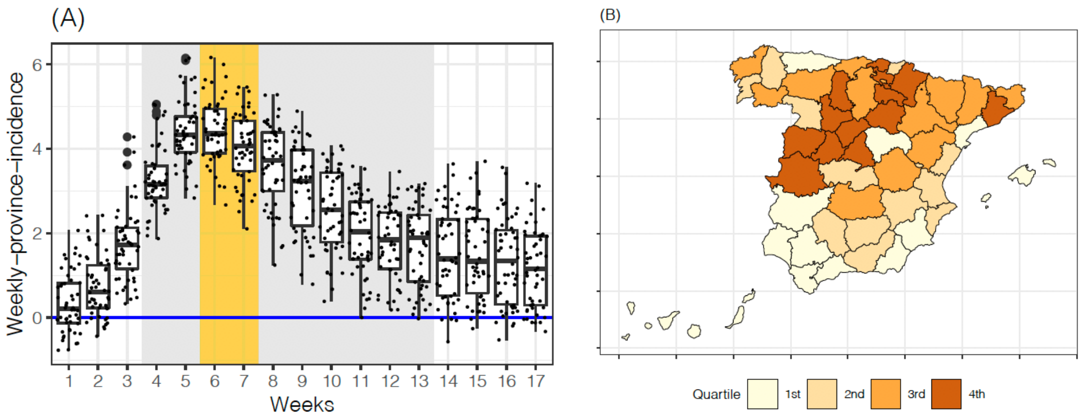

Figure 2A represents the weekly incidence of the pandemic (log of cases detected per 100,000 inhabitants). The incidence has an inverted-V shape that peaks on week 6. Figure 2B shows the spatial distribution of the incidence in week 6. A strong spatial pattern emerges, confirming the inter-provincial diffusion of the pandemic. Looking at these Figures, it is apparent that the incidence of COVID-19 has a strong impact on mobility in that movement decreases considerably as the virus intensifies.

The SUR approach

To begin with, we estimate a basic SUR model with G = 2 equations (for inter and intra movements, respectively), N = 50 individuals and Tm = 17 cross sections such as the following:

$$Within_{i,t} = \beta_{10} + \beta_{11} Emergence_t + \beta_{12} EmergenceTotal_t + \beta_{13} Density_{i,t} + \beta_{14} Old65_{i,t} + \beta_{15} Essential_{i,t} + \beta_{16} Incidence_{i,t-1} + \epsilon^W_{i,t}$$

$$Exits_{i,t} = \beta_{20} + \beta_{21} Emergence_t + \beta_{22} EmergenceTotal_t + \beta_{23} Density_{i,t} + \beta_{24} Old65_{i,t} + \beta_{25} Essential_{i,t} + \beta_{26} Incidence_{i,t-1} + \epsilon^E_{i,t}$$

$$cor(\epsilon^W_{i,t}, \epsilon^E_{j,t}) \;=\; \sigma_{i,j} \;;\quad t = 1, ..., 17; \quad i = 1, ..., 50$$

Within_{i,t} and Exits_{i,t} are respectively the mobility indices of the number of intra- and inter-provincial movements. A full description of all the variables is available in the help ?spsur::spain.covid of the spsur package.

The baseline model is the SUR-SIM specification of (1), which does not include spatial effects. To estimate this model, first, the two equations of model (1) can be specified using the \pkg{Formula} package. In this case, only a single sequence of explanatory variables is included on the right-hand side of the formula because these factors are the same for both equations.

```{r}
formula <- Within | Exits ~ Emergence + EmergenceTotal + Density +
  Old65 + Essential + Incidence
```

This specification can be estimated for maximum likelihood with the function spsurml

```{r}
Tm <- 17
covid.sim <- spsurml(formula = formula, data = spain.covid, type = "sim",
                   Tm = Tm, control = list(trace = FALSE))
summary(covid.sim)
```

```
Call:
spsurml(formula = formula, data = spain.covid, type = "sim",
    Tm = Tm, control = list(trace = FALSE))


Spatial SUR model type:  sim

Equation  1
                    Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)_1     0.75220735  0.04712812  15.9609 < 2.2e-16 ***
Emergence_1      -0.24061703  0.00593652 -40.5316 < 2.2e-16 ***
EmergenceTotal_1 -0.05211290  0.00795879  -6.5478 7.731e-11 ***
Density_1        -0.04233591  0.01700889  -2.4890  0.012905 *
Old65_1           0.34461164  0.06352238   5.4250 6.636e-08 ***
Essential_1       0.00177808  0.00066556   2.6716  0.007623 **
Incidence_1      -0.03576960  0.00219800 -16.2737 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
R-squared: 0.8643
  Equation  2
                   Estimate Std. Error  t value  Pr(>|t|)
(Intercept)_2     0.4349503  0.1135108   3.8318 0.0001319 ***
Emergence_2      -0.3132801  0.0142985 -21.9101 < 2.2e-16 ***
EmergenceTotal_2 -0.0585423  0.0191692  -3.0540 0.0022937 **
Density_2        -0.0513039  0.0409669  -1.2523 0.2106254
Old65_2           0.6620623  0.1529974   4.3273 1.598e-05 ***
Essential_2       0.0035088  0.0016030   2.1888 0.0287457 *
Incidence_2      -0.0405096  0.0052940  -7.6519 3.312e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
R-squared: 0.6338

Variance-Covariance Matrix of inter-equation residuals:
 0.004261810 0.007104235
 0.007104235 0.024723421
```

The residuals of both equations are highly correlated (0.69) and the Breush-Pagan test confirms the adequacy of the SUR specification. All the variables, with the exception of Density in the second equation, are significant. The signs of the two dummy variables, Emergence and EmergenceTotal are negative, as expected. The percentage of essential services, Essential, has a positive impact on movement in the sense that greater economic activity promotes mobility. Density comes with an unexpected negative sign. However, Density can be used as a pure measure of volume or as a fear factor, in the sense that the more people are concentrated, the higher the probability of contagion and the lesser the willingness to move. In our case, the last tendency prevails. The variable for older people, $Old65$, contrary to expectations, has a positive sign, which is difficult to justify. Finally, a strong incidence of the disease in previous weeks, Incidence_{t-1}, has a negative sign, which encourages the 'stay-at-home' policy by reducing both intra and inter-provincial movement.

Given the framework that supports our research and using Spanish provinces as observation units, it is advisable to test for the presence of (omitted) spatial effects, especially spatial autocorrelation. First, we need a weighting matrix that will be built according to the queen or simple contiguity criterion. The spdep package can be used at this moment. Note that there are three provinces with no neighbors (the Balearic and Canary islands), so the option zero.policy=TRUE will be necessary, as appears in the code below.

````
```{r}
listw <- spdep::poly2nb(spain.covid.sf, queen = TRUE)
listw <- spdep::nb2listw(listw, style = "W", zero.policy = TRUE)
```
````

The spsur function lmtestspsur gives assistance in testing the SUR-SIM model against any other specification through the corresponding Lagrange Multiplier. The results, as shown below, detect a strong spatial structure that has been omitted in the SUR-SIM: all the LM tests, robust and non-robust, reject their corresponding null hypotheses.

````
```{r}
covid.lmtest <- lmtestspsur(formula = formula, data = spain.covid,
Tm = Tm, listw = listw, zero.policy = TRUE)
pr.covid.lmtest <- sapply(covid.lmtest, broom::tidy)
print(as.data.frame(t(pr.covid.lmtest)))
```
````

Spatial SUR models. Selection strategy.

The next question is obvious: if the SUR-SIM model is not appropriate, which is the best specification, from a spatial perspective, to model the relation between mobility and incidence of COVID-19? Given our rather limited knowledge about this dilemma, we follow Lopez2014 by applying a formal model selection strategy based on a Specific-to-General, stge, combined with a General-to-Specific, gets, approach. Figure 3 shows the nesting structure that prevails among our eight spatial models, where the model to the left of the arrow is nested in the model to the right of the arrow. The difference between stge and gets is in the way of reading the chain: stge works from left to right, stopping the process when the model on the right is rejected in favor of the model on the left, whereas gets runs in the opposite direction.

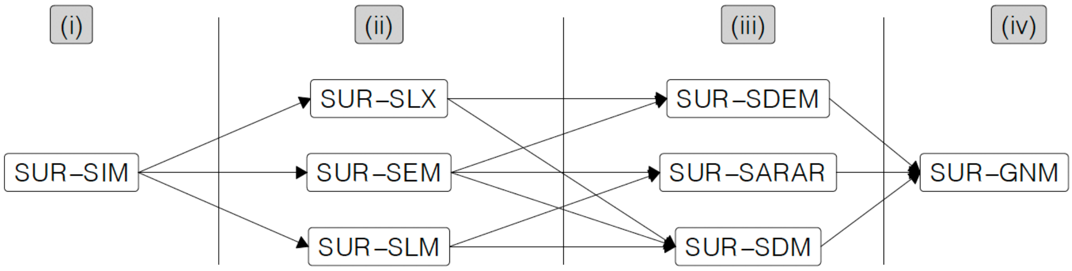

Figure 6: Nesting structure for the spatial models

Keywords: *spatial seemingly unrelated regression models, COVID-19, mobility, spsur R-package*
JEL codes: R-27, R-48