Territorial Patterns in COVID-19 Mortality in Spain, France, Italy and United Kingdom

Roberto Zavatta and Carlotta Moiso

5th Debate COVID-19, Forecast and Prediction
September 18th - 19th, 2020

ECONOMISTI ASSOCIATI

RATIONALE & SCOPE

Rationale

It all started in mid-March, when Italy was in the headlines worldwide for high COVID-19 mortality ..

.. yet, we knew that the bulk of deaths were concentrated in few areas.

Started looking around and noticed that the situation in other countries was similar, with significant differences across regions/provinces

Territorial differences progressively acknowledged, but media, researchers, modelers, etc. still largely focus at 'country level' ..

.. Some studies do investigate territorial differences, but typically one country at a time (e.g. Ehlert [2020] on Germany)

U.S.News & WORLD REPORT

No Let-Up in Coronavirus Deaths in Italy, New Cases Steady

Hospital in Alzano Lombardo, an area with extremely high mortality

Johns Hopkins has subnational data only for the US; France's INED only focuses on 'cross-country differences'; only OECD has a webpage on territorial differences

Scope

- ✓ **Focus on mortality (data less uncertain than cases…)**
- ✓ **Over the March – May period (i.e. the 'first wave')**
- ✓ **In Italy, France, Spain and UK (the most affected in Europe)**
- ✓ **At the sub-regional level (i.e. NUTS3 in EU jargon)**

NUTS3

Classification of territorial entities established by the EU in the 1970s. Includes macro-regions (NUTS1), such as Grand Est or England's North West, regions (NUTS2), such as Lorraine or Merseyside; and sub-regional entities (NUTS3), such as Moselle or Wirral. Focus on NUTS3 level, because data shows significant variations within NUTS2 (e.g. between Cáceres and Badajoz in Extremadura). Overall, the exercise covers 427 NUTS with some 143,000 COVID-19 deaths. Outermost regions excluded from the analysis due to lack of territorial contiguity.

ISSUES AND LIMITATIONS

Data on Deaths Not Fully Homogeneous

- ✓ **'Confirmed' deaths in Italy, France and Spain vs. 'certified' deaths in the UK.**
- ✓ **Tendency to underestimate deaths at home (possibly less so in the UK).**
- ✓ **For France and UK, data originate from different sources. For France and Spain some extrapolation/interpolation to fix the gaps.**
- ➡ **Unclear if this introduces any major bias across countries (would tend to say no)**

NUTS3 Show Differences Across Countries

- ✓ **Italy's *province* and France's *départements* are well established entities, broadly comparable in size and population (some 600,000 inhabitants on average).**
- ✓ **Spain's *provincias* are also historical entities, but significantly bigger and more populated than in other countries (about 900,000 people).**
- ✓ **In the UK, NUTS3 are a mix of historical counties (e.g. Dorset) and more or less 'artificial' groupings of local authorities (councils, etc.), and they are smaller (less than 400,000 people, on average).**
- ➡ **These differences do impact on the analysis and influence the ability to gather data**

Health Systems in the Four Countries Are Different

- ✓ **From near full centralization (France) to varying degrees of regionalization (Italy and Spain) to four separate systems (UK).**
- ➡ **Differences in the ability to collect sufficiently granular data**

TERRITORIAL CONCENTRATION OF DEATHS

High Concentration of COVID-19 Deaths

- ✓ Overall, 24% of deaths in just 10 NUTS3 areas accounting for only 11% of population.
- ✓ These include metropolitan areas (Paris and *petite couronne*, Madrid, Barcelona, Milan) but also medium-sized areas (Bergamo, Brescia, Haut-Rhin).
- ✓ Some large metropolitan areas (Roma, Naples, Bordeaux, Lyon, Marseille, Sevilla) were largely untouched.

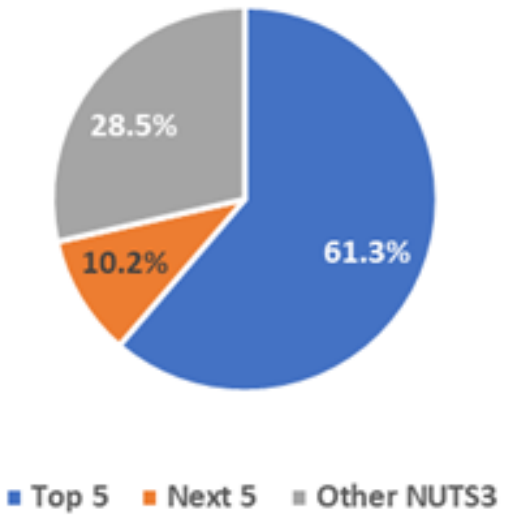
Major Differences Across Countries

- ✓ Concentration highest in Spain and lowest in the UK.
- ✓ In France, Italy and Spain, the share of deaths accounted for by the top 5 NUTS3 is two to three times the share of population.
- ✓ Much more uniform distribution in the UK (Top5: 9% of deaths and 7% of population)

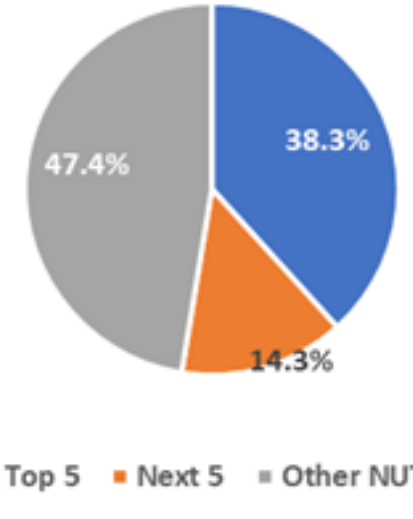
Important

- ✓ Results partly due to 'structural' differences across NUTS3 (Spain's are bigger, and obviously each one accounts for more deaths).

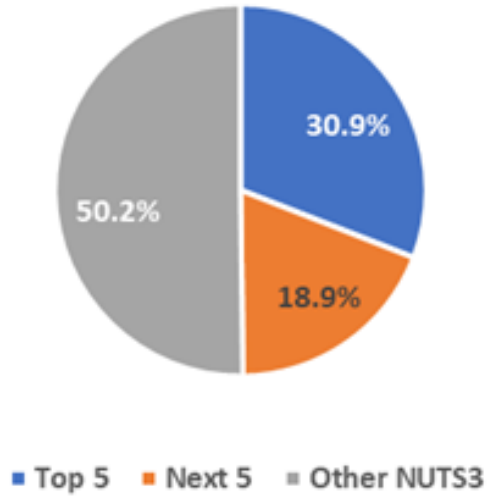
Spain



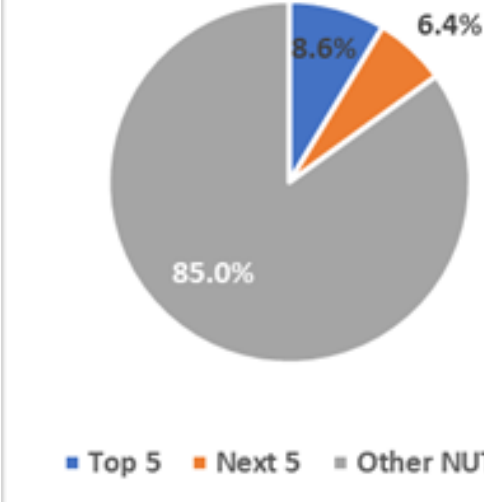
Italy



France



UK



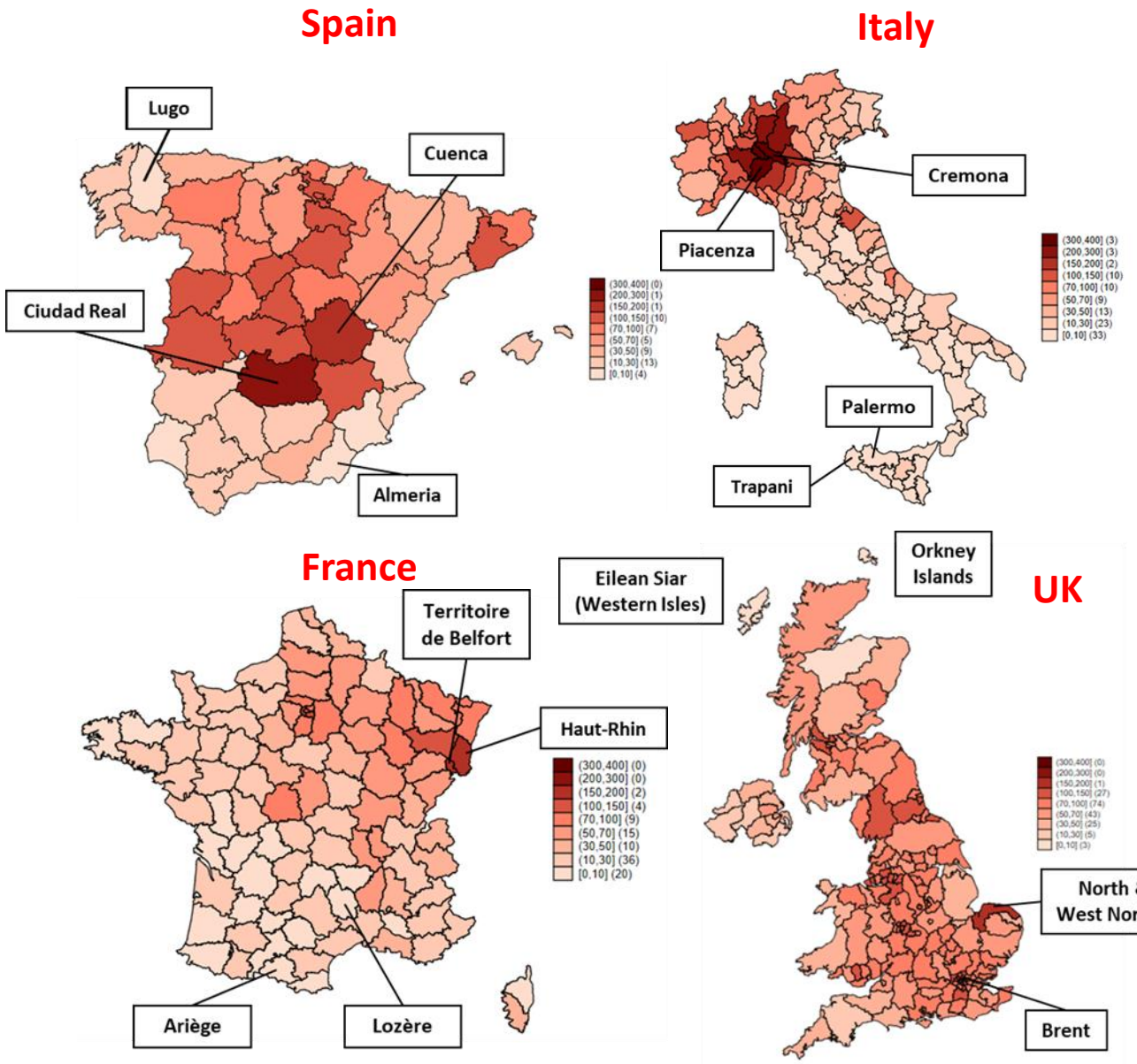
TERRITORIAL PATTERNS IN MORTALITY RATES (1)

Major Differences in Mortality Rates

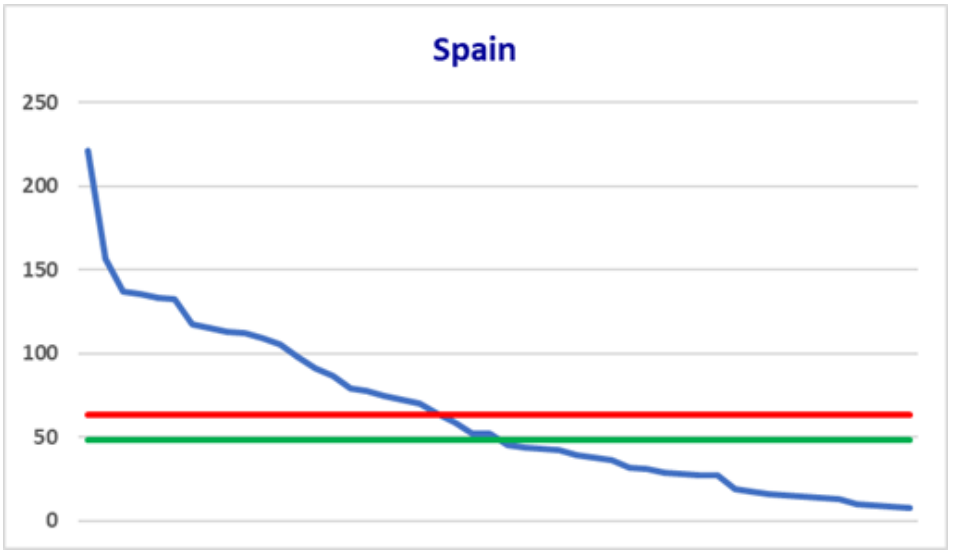
- ✓ Three NUTS3 areas with more then 300 deaths per 100,000, and another 4 with more than 200 per 100,000 ..
- ✓ .. but no less than 195 NUTS3 areas (i.e. nearly half of the total) with 50 deaths per 100,00 or fewer

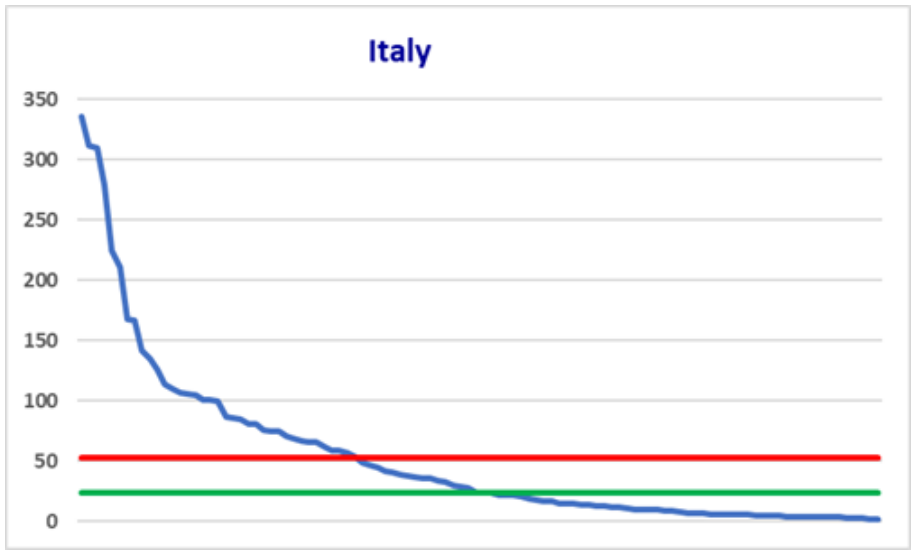
Once Again, Major Differences Across Countries

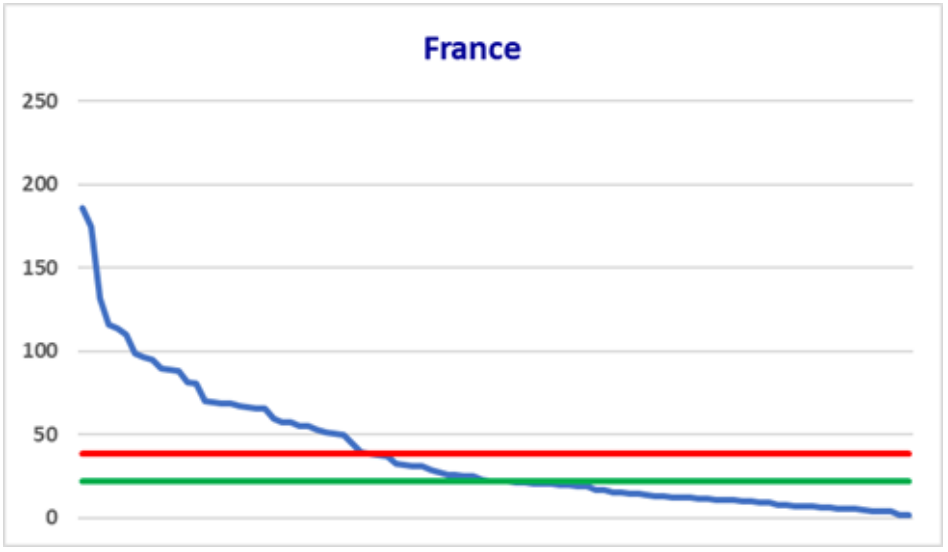
- ✓ Extremely polarized situation in Italy, where rates in the 5 most affected NUTS3 (in Lombardy & Emilia Romagna) are more than one hundred times those in the 5 least affected areas (mostly in Sicily). Mean 53, median 24
- ✓ Significant differences also in France, where rates in most affected areas (Grand Est & Paris) are 20 to 40 times bigger than in least affected (South West). Mean 38, median 22
- ✓ A much more uniform pattern is found in the UK, where a part a couple of peak areas there is a smooth, decreasing trend. Median is greater than the mean (76 vs. 74)
- ✓ Spain is somewhere in between. Mean 63, median 48.

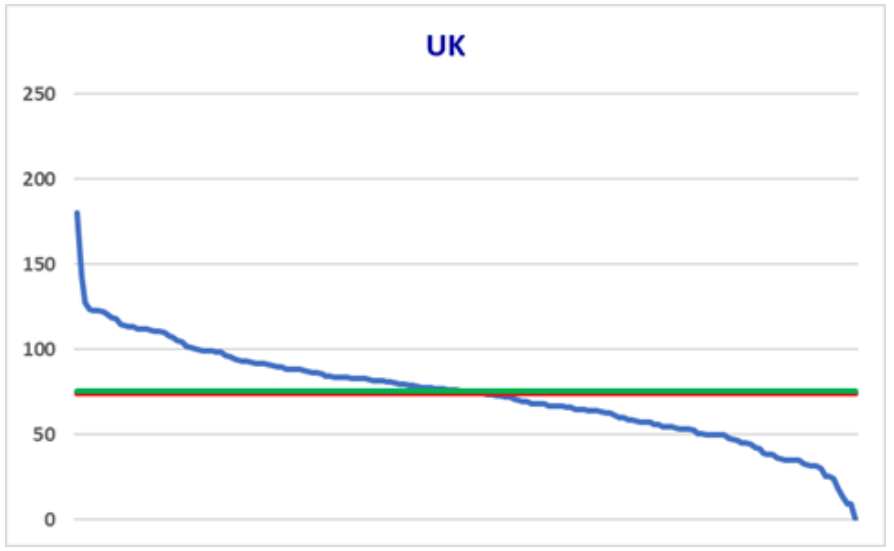


TERRITORIAL PATTERNS IN MORTALITY RATES (2)









Average in Red

Median in Green

Scale of vertical axis

- ✓ **0 – 350 for Italy**
- ✓ **0 – 250 for other countries**

DETERMINANTS OF TERRITORIAL DIFFERENCES (1)

Five categories of possible explanatory variables considered, namely:

- ✓ **Demographic factors;**
- ✓ **Socio-economic factors;**
- ✓ **Environmental and atmospheric conditions;**
- ✓ **Spatial effects;**
- ✓ **Health sector's features and policy response**

Combination of 'structural' and 'short-term' factors. In several cases, serious data limitations and/or mismatch between unit of analysis and explanatory variable.

Demographic Factors

- ✓ **Population density often mentioned as a facilitating factor, although this interplays with several other factors (see, for instance, Hamidi and others [2020]).**
- ✓ **Population age also frequently mentioned. It is regarded as a marker of the gradual accumulation of permanent damage over the life course (Dowd and others [2020]) and therefore is highly correlated with chronic diseases and conditions. In turn, there is ample evidence that preexisting chronic conditions are highly correlated with COVID-19 mortality**
- ✓ **Gender also considered in some studies, but with no meaningful results (Gujon and others [2020]).**

Socio-economic Factors

- ✓ **A priori, ambiguous relationship with income level. On the one hand, high income (GDP) is typically associated with travel, mobility, intense person-to-person interactions and therefore should be positively correlated, but …**
- ✓ **.. on the other, pandemics often affect the poor. Indeed, UK studies suggest deprivation conditions play a major role (ONS [2020]), but difficult to generalize due to lack of comparable data in other countries**

DETERMINANTS OF TERRITORIAL DIFFERENCES (2)

Environmental and Atmospheric Conditions

- ✓ It's a respiratory disease, so poor air quality intuitively associated with high mortality. Indeed, the impact of environmental conditions has been analyzed in several studies, although results are far from homogeneous (see, Caserini and others [2020] and Conticini and others [2020]). Also, sometimes there are considerable measurement issues (small number of observations for relatively large NUTS3 area, which affects representativeness .
- ✓ Temperature and other meteorologic conditions, such as humidity, considered by some studies (e.g. Paez and others [2020] and Briz-Redón A and Serrano-Aroca [2020]), but sometime with uncertain results.

Spatial Effects

- ✓ Linked to Tobler's first law, i.e. that "*All things are related, but nearby things are more related than distant things*". Various approaches can be considered. For the time being, we opted for as simple approach, i.e. the distance from the initial pandemic epicenter (implicitly assumes a sort of linear progression, which may be an oversimplification).
- ✓ Identification of initial epicenter not always straightforward. In Italy, considered Lodi. In France, Mulhouse (but Paris could also have been considered). In Spain, Madrid (but some also consider Euzkadi). In the UK, London was selected.

Health Sector's Features and Policy Response

- ✓ A priori, potentially very relevant, but limited information. Only the number of hospital beds is available across the four countries at the NUTS3 level (and for the UK a major effort to compile data was required)
- ✓ Mortality due to nosocomial infections could be proxied by data on health personnel infected, but data only for Italy and at NUTS2 level. Granular information on number of PCR tests also not available (only for Italy and Spain).

VARIABLES AND SOURCES OF DATA

Dependent Variable

- ✓ Mortality rate (deaths per 100,000, log) from various national sources.
- ➡ Important: for Spain, information was mostly taken from the dataset built by Estudio Montera, whose kind cooperation is gratefully acknowledged. The Estudio Montera's dataset is freely accessible at https://github.com/montera34/escovid19data

Explanatory Variables

- ✓ Average GDP per capita (log), from EUROSTAT
- ✓ Population density (inhabitants per square kilometer, log) from EUROSTAT
- ✓ Share of population aged 65 or more (percentage), from EUROSTAT
- ✓ Relative distance from epicenter (distance in kilometers/maximum distance within the country) own calculations
- ✓ Air quality (PM10 average annual concentration in 2018), from EEA and national sources
- ✓ Temperature (average temperature over the January – April 2020 period), from EU-funded Copernicus
- ✓ Hospital beds intensity (number of beds per 10,000), from national sources (with some approximation in the case of the UK)
- ✓ Metropolitan status (binary), based on European Commission classification of NUTS3
- ✓ Fixed country effects (binary)

RESULTS (PRELIMINARY, WORK IN PROGRESS)

VARIABLES	(1) ldrate	(2) ldrate	(3) ldrate	(4) ldrate	(5) ldrate	(6) ldrate	(7) ldrate
lgdp	0.666***	0.687***	0.133	0.0729	0.109	-0.0627	-0.00257
	(0.209)	(0.209)	(0.141)	(0.134)	(0.140)	(0.113)	(0.119)
lpopdens	0.151***	0.178***	0.0101	0.0210	0.00115	0.0378	0.0932**
	(0.0439)	(0.0436)	(0.0441)	(0.0390)	(0.0458)	(0.0368)	(0.0448)
old	-0.137	4.123***	0.369	-1.663	-0.254	-1.688	1.201
	(1.325)	(1.366)	(1.146)	(1.101)	(1.169)	(1.064)	(1.149)
distance_norm1			-3.022***	-3.358***	-3.042***	-3.025***	-2.541***
			(0.219)	(0.210)	(0.217)	(0.213)	(0.231)
group(countrycode) = 2, FR		-0.935***	-0.421***		-0.676***		-0.705***
		(0.170)	(0.134)		(0.177)		(0.166)
group(countrycode) = 3, IT		-0.971***	-0.428***		-0.383**		-0.799***
		(0.190)	(0.153)		(0.161)		(0.167)
group(countrycode) = 4, UK		-0.167	0.133		0.219		-0.195
		(0.148)	(0.138)		(0.155)		(0.155)
met	0.113	0.0333	0.111	0.141*	0.0885	0.177**	0.0774
	(0.107)	(0.106)	(0.0819)	(0.0801)	(0.0819)	(0.0739)	(0.0740)
level_PM10				-0.0242***	0.00495	-0.0102	0.0179*
				(0.00796)	(0.0101)	(0.00777)	(0.0105)
beds_10000				-0.00323	0.0103*	-0.00296	0.00351
				(0.00278)	(0.00557)	(0.00302)	(0.00570)
temperature						-0.112***	-0.131***
						(0.0211)	(0.0228)
Constant	-3.962*	-4.668**	3.157**	4.586***	3.143**	6.407***	4.687***
	(2.164)	(2.156)	(1.547)	(1.427)	(1.528)	(1.250)	(1.318)
Observations	427	427	427	427	427	427	427
R-squared	0.191	0.324	0.567	0.535	0.575	0.573	0.614

Four Countries Cumulated

*** p<0.01, ** p<0.05, * p<0.1

p<0.01 highlighted in green

p<0.05 highlighted in light blue

Four Countries Individually

Distance from epicenter highly significant for all countries (but lower coefficient in UK)

Air pollution significant for Italy, temperature mildly significant in all countries

Socio-economic and demographic features play a limited role, except in UK

Hospital beds intensity significant only in the UK

R-squared is around .6/.7 in Italy, France and Spain and about .5 in the UK

CONCLUSIONS & NEXT STEPS

Interpreting Results

- ✓ **Why is the distance from epicenter so important, especially in Italy, France and Spain? Does it make any sense? In our view yes, due to two factors**
  - ❑ **First, it captures the 'lockdown effect', which interrupted/greatly slowed down territorial transmission. And the earlier the adoption of the lockdown the stronger the effect. Hence difference in coefficients between Italy, France and Spain (lockdowns adopted in early/mid March) and the UK (lockdown enforced later in March)**
  - ❑ **Second, it may also capture the occurrence of major outbreaks in hospitals and nursing homes, which mostly (although not exclusively) occurred in the epicenter and neighboring areas, due to unpreparedness of health care structures (missing granular data, but plenty of 'qualitative' evidence in this respect**
- ✓ **In the UK, distance from epicenter is significant but less important and other factors play a greater role (especially, demographic ones)**

Next Steps

- ✓ **Refine the analysis and add the econometric exercise to the initial descriptive paper published in medRxiv (https://www.medrxiv.org/content/10.1101/2020.07.27.20162677v1)**
- ✓ **May extend the analysis to other countries**
- ✓ **More than happy to cooperate with other researchers and willing to share our highly granular dataset**

THANK YOU!!

ANNEX TABLES (1)

Territorial Concentration of COVID-19 Deaths

Spain			France			Italy			UK		
Top 10 NUTS3	Share of Deaths	Share of Population	Top 10 NUTS3	Share of Deaths	Share of Population	Top 10 NUTS3	Share of Deaths	Share of Population	Top 10 NUTS3	Share of Deaths	Share of Population
Madrid (Madrid)	30.7%	14.9%	Paris (Île de France)	8.1%	3.3%	Milano (Lombardia)	12.7%	5.4%	Birmingham (West Midlands)	2.3%	1.7%
Barcelona (Catalunya)	21.2%	12.5%	Hauts-de-Seine (Île de France)	6.4%	2.5%	Bergamo (Lombardia)	9.6%	1.8%	Hertfordshire (East of England)	2.0%	1.8%
Ciudad Real (Castilla-la Mancha)	3.7%	1.1%	Val-de-Marne (Île de France)	6.3%	2.2%	Brescia (Lombardia)	8.3%	2.1%	Staffordshire (West Midlands)	1.5%	1.3%
Bizkaia (País Vasco)	3.0%	2.6%	Seine-Saint-Denis (Île de France)	5.4%	2.6%	Torino (Piemonte)	4.0%	3.7%	Tyneside (North East)	1.5%	1.3%
Toledo (Castilla-la Mancha)	2.6%	1.6%	Haut-Rhin (Grand Est)	4.8%	1.2%	Pavia (Lombardia)	3.8%	0.9%	Berkshire (South East)	1.4%	1.4%
Valencia (Comunidad Valenciana)	2.4%	5.7%	Val-d'Oise (Île de France)	4.1%	1.9%	Cremona (Lombardia)	3.5%	0.6%	Harrow & Hillingdon (London)	1.3%	0.8%
Zaragoza (Aragón)	2.3%	2.2%	Seine-et-Marne (Île de France)	3.9%	2.2%	Piacenza (Emilia-Romagna)	3.0%	0.5%	Barnsley, Doncaster and Rotherham (Yorkshire & Humber)	1.3%	1.2%
Girona (Catalunya)	1.9%	1.7%	Rhône (Auvergne-Rhône-Alpes)	3.8%	2.9%	Genova (Liguria)	2.7%	1.4%	Leeds (Yorkshire & Humber)	1.3%	1.2%
Albacete (Castilla-la Mancha)	1.8%	0.9%	Essonne (Île de France)	3.6%	2.0%	Monza (Lombardia)	2.7%	1.4%	Greater Manchester NE (North West)	1.2%	1.0%
Navarra (Navarra)	1.7%	1.5%	Moselle (Grand Est)	3.5%	1.6%	Parma (Emilia-Romagna)	2.3%	0.7%	Durham (North East)	1.2%	0.8%
Concentration Ratios			**Concentration Ratios**			**Concentration Ratios**			**Concentration Ratios**		
Top 5	61.3%	32.6%	Top 5	31.0%	11.7%	Top 5	38.3%	14.0%	Top 5	8.6%	7.5%
Next 5	10.2%	11.9%	Next 5	18.9%	10.6%	Next 5	14.3%	4.7%	Next 5	6.4%	5.0%
Top 10	71.5%	44.5%	Top 10	49.8%	22.3%	Top 10	52.6%	18.6%	Top 10	15.0%	12.5%

ANNEX TABLES (2)

Territorial Distribution of Mortality Rates (COVID-19 Deaths per 100,000)

France		Italy		Spain		UK	
Top 5 NUTS3	**Rates**	**Top 5 NUTS3**	**Rates**	**Top 5 NUTS3**	**Rates**	**Top 5 NUTS3**	**Rates**
Haut-Rhin (Grand Est)	185.8	Piacenza (Emilia-Romagna)	335.7	Ciudad Real (Castilla-la Mancha)	221.4	North & West Norfolk (East of England)	180.6
Territoire de Belfort (Bourgogne-Franche-Comté)	174.4	Cremona (Lombardia)	311.5	Cuenca (Castilla-la Mancha)	156.8	Brent (London)	143.2
Val-de-Marne (Île de France)	131.5	Lodi (Lombardia)	309.3	Madrid	136.6	Walsall (West Midlands)	127.6
Hauts-de-Seine (Île de France)	116.0	Bergamo (Lombardia)	279.0	Albacete (Castilla-la Mancha)	135.5	Harrow & Hillingdon (London)	123.2
Vosges (Grand Est)	114.0	Pavia (Lombardia)	224.4	Soria (Castilla y León)	133.0	Wirral (North West)	123.1
Bottom 5 NUTS3	**Rates**	**Bottom 5 NUTS3**	**Rates**	**Bottom 5 NUTS3**	**Rates**	**Bottom 5 NUTS3**	**Rates**
Dordogne (Nouvelle-Aquitaine)	4.1	L'Aquila (Abruzzo)	3.0	Cádiz (Andalucía)	12.6	Derry City and Strabane (Northern Ireland)	17.9
Landes (Nouvelle-Aquitaine)	3.9	Oristano (Sardegna)	2.5	Murcia (Murcia)	9.8	Fermanagh and Omagh (Northern Ireland)	13.7
Lot-et-Garonne (Nouvelle-Aquitaine)	3.6	Ragusa (Sicilia)	2.2	Huelva (Andalucía)	9.2	Inverness & Nairn and Moray, Badenoch & Strathspey (Scotland)	9.3
Ariège (Occitanie)	1.3	Trapani (Sicilia)	1.6	Lugo (Galicia)	8.2	Orkney Islands (Scotland)	9.1
Lozère (Occitanie)	1.3	Palermo (Sicilia)	1.4	Almería (Andalucía)	7.6	Western Isles (Scotland)	0.0
Summary Statistics	**Rates**	**Summary Statistics**	**Rates**	**Summary Statistics**	**Rates**	**Summary Statistics**	**Rates**
Average	38.2	Average	46.1	Average	63.2	Average	74.0
Median	21.8	Median	20.0	Median	48.2	Median	75.6
Coefficient of Variation	98%	Coefficient of Variation	137%	Coefficient of Variation	76%	Coefficient of Variation	37%

SELECTED BIBLIOGRAPHY

- ✓ Briz-Redón A and Serrano-Aroca A, A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain, Science of the Total Environment, April 2020
- ✓ Caserini and others, Pollution and Covid. Two vague clues don't make an evidence, Scienza in rete, 28 April 2020
- ✓ Conticini E and others, Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environmental Pollution, April 2020
- ✓ Dowd J B and others, Demographic science aids in understanding the spread and fatality rates of COVID-19 PNAS May 5, 2020
- ✓ Ehlert A, The socio-economic determinants of COVID-19: A spatial analysis of German county level data, June 25, 2020
- ✓ Goujon A and others, Age, gender, and territory of COVID-19 infections and fatalities, JRC Technical Report, 2020
- ✓ Hamidi S and others, Does Density Aggravate the COVID-19 Pandemic? Early Findings and Lessons for Planners, Journal of the American Planning Association, 2020
- ✓ Office of National Statistics (ONS), Deaths involving COVID19 by local area and deprivation, 1 May
- ✓ Paez A and others, A Spatio-Temporal Analysis of the Environmental Correlates of COVID-19 Incidence in Spain, Geographical Analysis, May 2020