COVID, We Are in This Fight Together

David Wei, Imane Farhat, Pak Shing Ho, Shahbaz Mogal

Executive summary

COVID-19 has brought to the world as much division and mistrust as physical suffering and deaths. Countries have adopted the mindset to help themselves, while alienating working with countries claimed to be "epicenters" or those with policies different to their own. Within the same country itself, sections of the society have divided themselves into subgroups obeying rules that benefit them, while turning a blind eye to other rules at the expense of other subgroups of society. This study shows that moving away from COVID-19 will take a more concerted effort from all countries and all subgroups within countries, and this will be beneficial to all groups involved.

More specifically, we first explore the effects countries have on each other. The success of fighting COVID-19 in a specific country involves coordinating policies with neighboring countries. Moreover, we explore the effects different age groups within the same country have on each other, and how reducing the impact on a specific age group will require other age groups adhering to policies that apply to them as well. The following results support our conclusions:

- Countries that are **geographically close** to each other experience similar COVID-19 severity to COVID-19 waves in terms of correlation and lead-lag synchronicity (Fig. A).
- Responses of **different age groups** within a country are also closely related to each other.
- In most cases there are **time lags** in response between two sections, the time lags are very informative information on preventing the spread of the disease (Fig. B).
- The implication of our work includes:
 - Policy-makers should take the relationship of different sections into account.
 For example, how reopening its border affects the neighboring countries or how reopening schools affects the health well-being of the elders.
 - Countries should work together when implementing a policy in order to achieve the best effects.

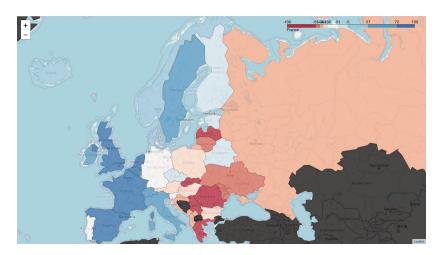


Figure A: **Correlation with France** in terms of new deaths per day. Neighboring countries of France show positive correlation (blue), distant countries in Eastern Europe show negative correlation (red) with France

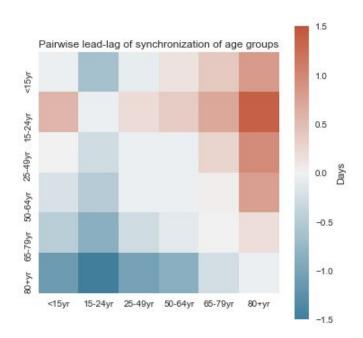


Figure B: Lead-lag of synchronization between weekly new cases of age groups. Roughly speaking, younger groups lead older groups by a week or so.

1. Introduction

Coronavirus disease 2019 (COVID-19) is a contagious respiratory disease. It was first discovered in Wuhan, China in December 2019 and quickly spread across the whole world in the next few months (Page et al., 2021). About 20% of patients experienced severe to critical conditions (Centers for Disease Control and Prevention, CDC, 2020). According to WHO (World Health Organization), COVID-19 has caused 116 million cases and 2.59 million deaths worldwide as of March 2021.

One of the biggest challenges in containing the spread of COVID-19 is from the division and mistrust of the public on government regulations and scientific guidance. For instance, A CDC survey shows that young adults are less likely to wear masks and take other measures against COVID-19, possibly because they are "less likely to suffer the severe illness from COVID-19" (Hutchins et al., 2020). For another example, former President Donald Trump argued for premature school reopening because "not that many of you (students) are going to die". Such misinformed and misleading opinions have greatly facilitated the spread and damages of COVID-19 in many states and countries. A recent example would be that the Texan governor Greg Abbott is "reopening businesses at 100% capacity" despite the suggestions from health officials and the latest weekly average of 200 deaths per day in Texas.

The purpose of this study is threefold. First, we want to know how closely different sections of a society are connected together and influence each other in response to COVID-19. If all the sections show strong connection to one another, it is necessary for the entire society to put away the self-preservation instinct and join forces in fighting COVID-19. Second, we are also interested in the short term prediction of the outbreak of COVID-19. If we can statistically identify the speed and nature of the spread from one group to another, policy-makers can accordingly make responses in time. Lastly, it is also necessary to quantify the effectiveness of a policy in order to avoid unnecessary negative impact to people's daily lives and the economy.

Our work shows strong connections between the responses of different sections of the society, namely, the connection between neighboring countries and the connection between different age groups of a nation. We also identify and quantify the transmission processes of COVID-19 with composite analyses and cross-correlation analysis. The results show the

propagation of the COVID-19 cases from younger age groups to senior age groups on the time scale of a week, and from one country to another on the time scale from a week to two months, depending on the geographical distance. Lastly, we propose and validate a measure which quantifies the average effect of a policy. The rest of the report is arranged as follows. Section 2 introduces the data and methodology of this study. Section 3 analyzes and discusses the results. Finally, section 4 summarizes the conclusions and discusses their implications.

2. Data and Methods

The datasets used in this study are provided by the organizers. It was compiled from three sources, "Our World in Data", the "European Centre for Disease Prevention and Control" (ECDC) and the "Covid Tracking Project". Since our study focuses on the relationship in responses between different sections of a society, we mainly use the ECDC data to study the interaction between countries and the age-group data from ECDC to study the interaction between age groups.

Three main mathematical models are used in this study. First, we use composite analyses to quantify the average effect of a policy. Specifically, for a certain policy such as closing daycare centers, we collect all related statistics of such instances as well as their history ranging from five weeks before the implementation of the policy and five weeks after that. Then we make pointwise averages of the collection of these time series to obtain the composite mean, which is assumed to have averaged out the effects of all confounding variables.

Second, to quantify the effect of a policy, we devise a measure that calculates the second derivative of new COVID-19 case, which writes

$$\frac{new_cases(-5 weeks) + new_cases(+5 weeks) - 2 \times new_cases(now)}{(10 weeks)^2}$$

We examine all 60 policies with this measure and find 6 policies that are effective statistically.

Lastly, we also build a simple pipeline to investigate the contemporaneous correlation and lead-lag synchronicity of COVID-19 severity between age groups and between European

countries¹. The contemporaneous correlation is measured as Spearman's rank correlation of two series between countries after removing the corresponding average trend across all age groups/European countries. Without removing the cross-section time trend nor using a ranked correlation, the correlation could yield a result dominated by the large values from the two peaks and possible outliers. To put it in mathematical terms,

- Variable at day $t: x_{it}$.
- Removing aggregate average time trend (cross-section de-mean): $y_{it} = x_{it} \sum_i x_{it}$. In other words, we would like to distill the country-specific component by removing the Europe-specific component from the original series.

• Spearman's rank correlation:
$$r_{ij} = 1 - 6\sum_{t} \frac{rank(y_{it}) - rank(y_{jt})}{n(n^2 - 1)}$$

The lead-lag synchronicity is measured as the number of leads or lags yielding the highest Pearson's correlation of two series between countries. Specifically,

$$\bullet \quad \textit{offset days} = argmax_{\tau}r_{ij\tau} \text{ , where } r_{ij\tau} = \frac{cov(x_{it}, x_{jt-\tau})}{\textit{std}(x_{it})*\textit{std}(x_{jt-\tau})} \text{ , for } \tau \in [-60, \ 60] \text{ .}$$

3. Analyses and Results

Variations on a global scale: Geographical Correlation of COVID-19
 Impact

We study the correlation between the impact a country has on the countries that surround it. Our results match our hypothesis that countries close-by each other have a positive correlation in terms of COVID-19's effect on the countries, while countries far away have insignificant/negative correlation. As explained in our method section, we have distilled the country-specific component by removing the Europe-specific component from the original series, so that the correlation would not yield a result dominated by the large values from the two wave peaks and possible outliers. This possibly explains why distant countries are negatively correlated.

¹ Contemporaneous correlation and lead-lag synchronicity results of different series can be generated from our codes submitted.

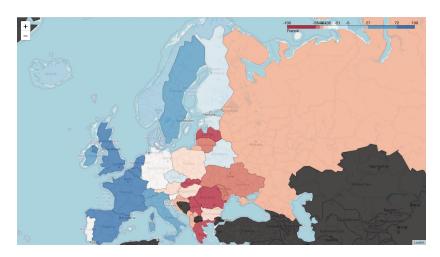
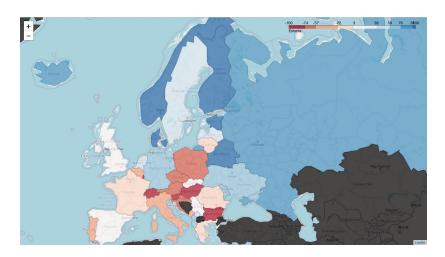


Figure above: **Correlation with France** in terms of new deaths per day. Neighboring countries of France show positive correlation (blue), distant countries in Eastern Europe show negative correlation (red) with France

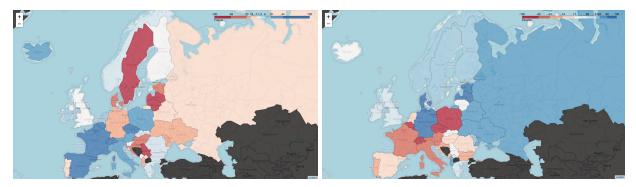


Eigure above: **Correlation with Estonia** in terms of new deaths per day. Neighboring countries of Estonia show positive correlation (blue), distant countries in Western Europe show negative correlation (red) with Estonia

We computed the pairwise correlation of the new deaths per day due to COVID-19 between countries across Europe. Two examples are presented above. We find that neighboring countries have synchronized COVID-19 impacts, despite varying policies and cultures. There are a few outliers, the most significant being Germany outperforming its neighbors. This is not surprising in that Germany did a much better job containing COVID-19.

We also find synchronization of COVID-19 death rates between clusters of countries as business-friendly blocks within Europe (Western Europe/Scandinavia/Eastern Europe) and

those within travelling distance of each other. This supports our hypothesis that in order to control the impact of COVID-19, countries also need to cooperate with their neighboring countries and coordinate their policies with each other.



Figures above: **Correlations with France (left) and Estonia (right)** in terms of the spread of new cases per day show synchronization between the corresponding neighboring countries.

We see that the spread of new COVID-19 cases per 100k between countries is also synchronized between neighbors. Estonia, for example, shows positive (blue) correlations with its neighbors (Latvia, Russia, Scandinavia) and negative (red) correlations with distant countries (France, Italy, Austria). This further supports our argument that simply focusing on a country's own policy is not doing enough to fight COVID-19. Countries should actively work with their neighbors in containing COVID-19 case spread, which directly leads to deaths, as examined before. Synchronization of the effect of COVID-19 also creates an opportunity for countries to learn from their neighbors' mistakes and improve their future policies.

3.2. Variations on a global scale: Prepare Early, Brace for the Impact

We have shown that neighboring countries exhibit a significant correlation between COVID-19 waves. To further our study into how countries can use this information to dampen the impact of COVID-19, we investigate how COVID-19 spreads across neighboring countries. Our most important finding in this was that there is a time lag between countries getting hit by COVID-19, and this time lag is related to the geographical distance.

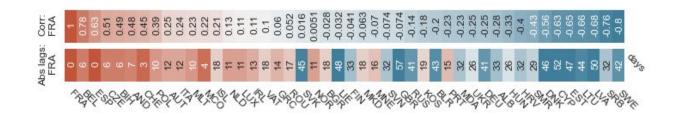


Figure above: **Correlation** and **Lead-Lag Synchronization** of Countries with France – Positive

Correlation and Lower Absolute Lead-Lag Days for Neighboring Countries, Negative Correlation and

Higher Absolute Lead-Lag Days for Distant Countries

As shown in the figure above, countries neighboring France (Belgium, Spain) have both more positive correlation to France, and higher degree of synchronization (0-7 days). Countries distant to France (Serbia, Sweden) have negative correlation and lower degree of synchronization (30-50 days). This is an important result as it indicates that COVID-19 slowly yet certainly traverses countries at the expected time, despite variation in policies and strategies. Fighting COVID-19 is not effective through the power of just one country for its own population without any time-sensitive coordination² across countries.

The relationship between the geographical distances and the synchronization of COVID-19 impact implies that countries can anticipate and prepare for the waves of COVID-19 approaching from their neighboring countries. Having an estimate of the timing of the COVID-19 propagation not only provides valuable preparation time but also provides insights on the international coordinations.

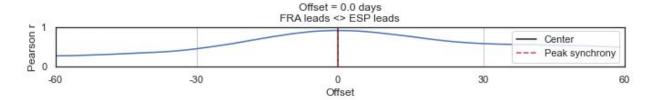


Figure above: New cases per million highly correlated and synchronized between Spain & France

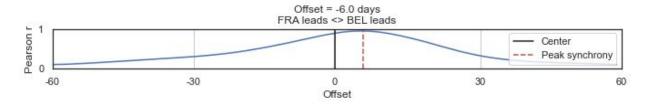


Figure above: Belgium leads France by 6 days in new cases per million

8

² In theory, an extreme example of perfect coordination would be quarantine at home for anybody on the planet for a period of time starting at the same time.

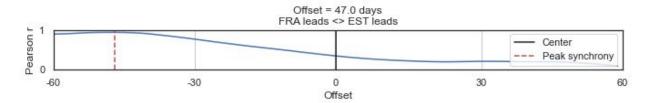


Figure above: France leads Estonia by more than a month in new cases per million

The figures above provide a more intuitive demonstration of the lead-lag of synchronization between countries. France and Spain, which border each other, are almost perfectly synchronized in their numbers of new COVID-19 cases. France lags Belgium by 6 days. Note that this is not due to a global simultaneous outbreak, as Estonia lags France by 47 days. Instead, this is likely due to the virus propagation via travel across neighboring borders. Preparing early for this expected propagation, bracing for the impact by building up hospital capabilities and implementing effective policies in advance would help reduce deaths, and slow down the propagation of the virus.

3.3. Variations within a Country: COVID-19 Impact on Sub-groups Requires Policies from a Holistic Perspective

Next, we investigate how countries not only need to coordinate on a global scale, but also domestically. To analyze the domestic transmission of COVID-19, we divide the population of a country into subgroups based on their age. We pick age as our choice of study because most countries choose age as the main aspect to design age-differentiating policies. Examples include lifting bans on daycare when new cases in children are low. Our results show that age-differentiating policies have a broader impact through transmission between sub-groups and should be evaluated from a holistic perspective.

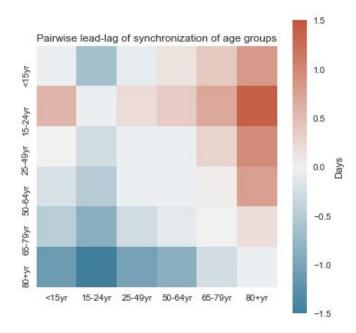


Figure above: Lead-lag of synchronization between weekly new cases of age groups. Roughly speaking, younger groups lead older groups by a week or so.

The figure above shows the pairwise lead-lag of synchronization between weekly new cases of age groups. In general, the upper right corner has positive values and the lower left corner has negative values, which indicates that younger groups tend to lead older groups by a week or so. The bigger the age difference between two groups, the longer is the lag. This consistent pattern of the lead-lag of synchronization likely implies an underlying causal relationship between the age groups -- an asymmetric transmission of COVID-19 cases from younger to older generations.

	ClosDaycare (new cases per 100k) Weeks										
	A	35	າ	>	0	^	2	ზ	٨	6	6
<15yr	99.025	126.44	124.329	105.537	41.5118	33.575	11.0435	11.4917	10.1542	9.4625	9.23043
15-24yr	462.35	516.74	478.586	423.125	162.6	119.55	51.8957	48.2875	41.0792	38.0708	37.5957
25-49уг	350.875	371.4	397.586	414.087	175.547	139.845	86.8043	77.9792	63.3292	58.0833	54.487
50-64yr	302.825	290.24	312.214	341.775	156.194	137.08	97.0913	93.2375	74.8125	64.3792	55.813
65-79yr	197.45	168.82	196.129	231.625	110.659	102.685	84.2826	83.675	65.4958	52.5458	44.2174
80+yr	314.1	242.1	369.157	480.15	222.735	194.165	153.235	167.088	162.475	165.525	167.439

Figure above: Composite mean of new cases per 100k across age groups when closing daycare centers at week 0. There is a uniform drop of new cases in week 0.

To further validate the broader impact of a policy, we perform the composite analysis on age-differentiating policies. The figure above shows a 10-week time series of new cases per 100k averaged across all instances of closing daycare centers in Europe. Even though closing daycare centers is only targeting younger age groups, the figure shows a uniform drop in new cases across all age groups at around week 0. This is likely due to the prevention of the one-way transmission from younger to older generations. Even an age-specific policy could effectively influence all sub-groups. Therefore, a policy-maker should avoid isolating the impact of a policy and evaluate a policy from a holistic perspective.

3.4. Quantifying the effect of a policy

In this section, we propose a measure to estimate the effect of a policy on the time scale of two and a half months, and validate all 60 policies with data. The measure is a classic three-point finite difference scheme of the second derivative of the new cases as follows,

$$\frac{new_cases(-5 weeks) + new_cases(+5 weeks) - 2 \times new_cases(now)}{(10 weeks)^2}$$

The physical meaning of this measure is the increase of increase in new cases. If the new cases increase linearly with time, the measure gives 0. If the new cases increase slower and slower with time, the measure would be negative. We apply this measure to all 60 policies from the dataset stored in "2_ecdc/country_response_measures", and find 6 statistically effective policies, namely, "ClosDaycare", "EntertainmentVenues", "ClosSec", "ClosHigh", "BanOnAllEvents" and "GymsSportsCentres". An example for "EntertainmentVenues" is shown below.

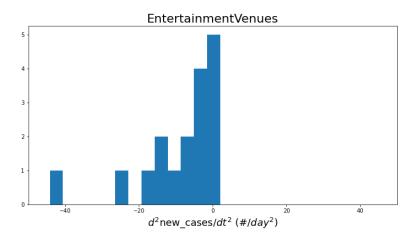


Figure above: histogram of the effect of shutting down entertainment venues. All cases show a negative second derivative (decrease in increase) of the new cases over the course of 10 weeks.

The figure above shows the histogram of the effect of shutting down all entertainment venues. There are a total of 16 cases of such policy and all of them decrease the increase of new cases per 100k. The result suggests an effective and consistent impact of shutting down all entertainment venues, and proves the reliability of measuring the effect of a policy with the finite-difference second derivatives.

4. Conclusions and Discussions

We have studied the interactions between the impact of COVID-19 on different sections, namely, countries and age groups. With the datasets provided by the organizer, and a mathematical framework involving composite analyses and cross correlation, we identified the

propagation of COVID-19 waves from one country to another, and from younger to older generations. We found that the transmission between age groups within a country takes about a week, while the transmission from one country to another takes from one week to two months, depending on the geographical distance. The further two age groups are apart, the longer the geographical distance, the longer it takes for the propagation of COVID-19 waves. The result suggests that an effective policy needs to take a holistic perspective for both consideration of coordination between countries and between domestic sub-groups. In addition, we proposed to use a second-order derivative of the finite difference scheme to quantify the effect of a policy. We validated this measure with the data and identified 6 effective policies.

The conclusions suggest that nobody is alone in the battle with COVID-19 --- we are all in this fight together! The safeness of a section is highly dependent on the other sections. As a result, the society needs to put away the pure self-preservation instinct and join forces in the fight. The enemies of our public enemy (virus) are indeed our friends! The policy-makers should not make isolated decisions based on a specific fraction of the society. The conclusions have many implications to the nowadays world. For example, reopening schools is not necessarily a good idea during the pandemic, considering the likelihood of the transmission from children to elders. Banning exports of medical supplies to neighboring countries risks suffering backfires in the long term.

5. Appendix

Gantt Chart of A Subsample of Response Measures

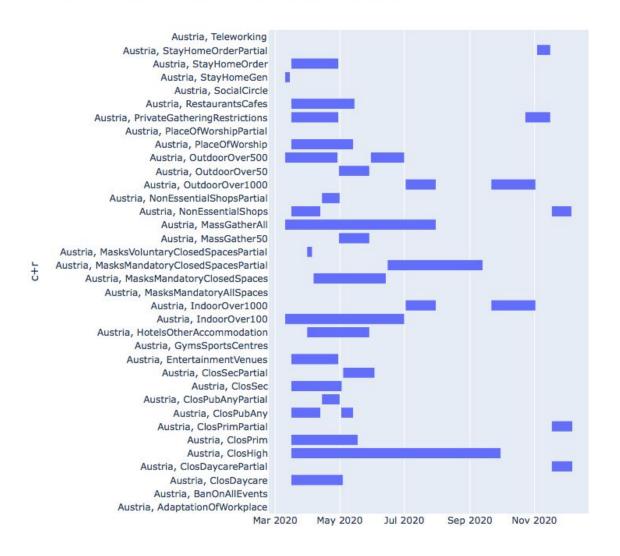
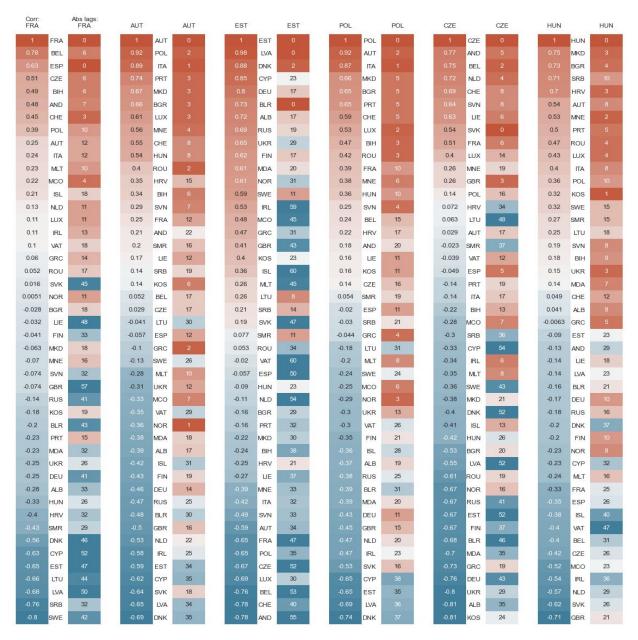
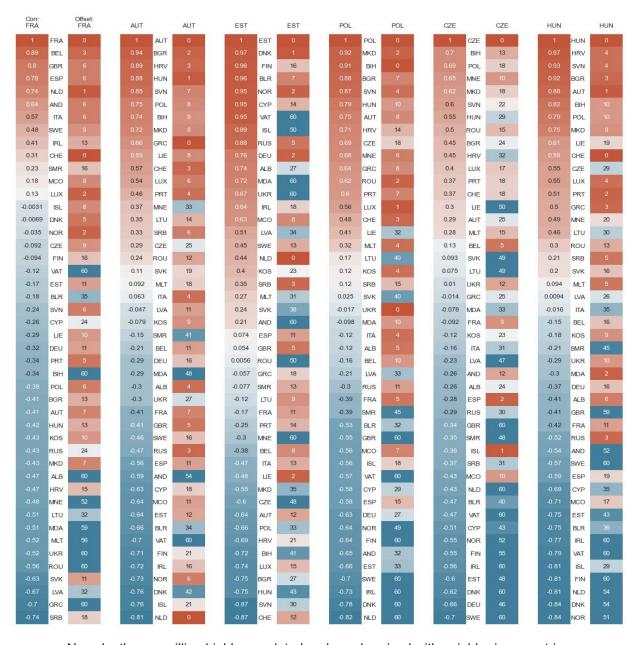


Figure above: Gantt chart of a subsample of the European response measures.

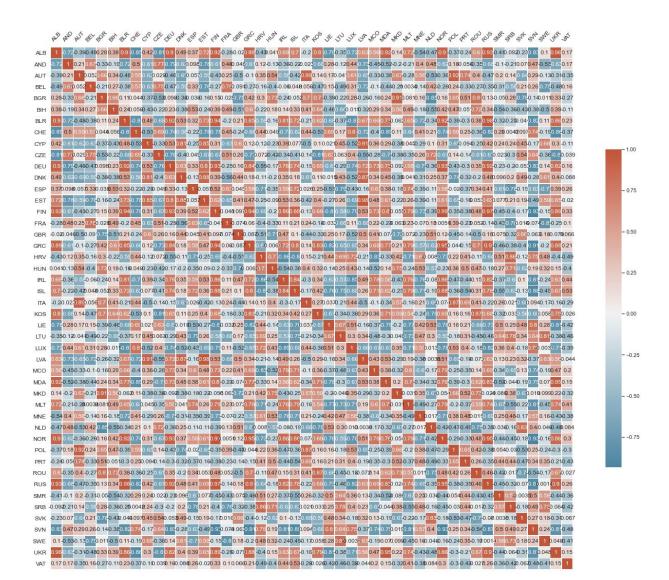


New cases per million highly correlated and synchronized with neighboring countries

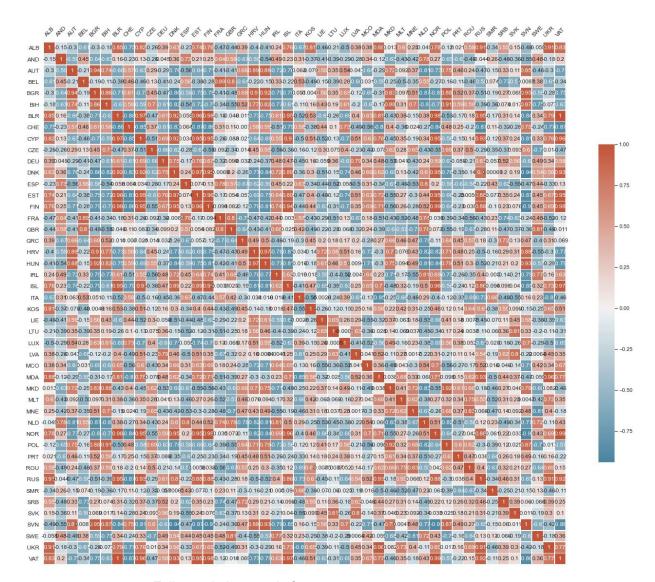
© 2021 David Wei, Imane Farhat, Pak Shing Ho, Shahbaz Mogal



New deaths per million highly correlated and synchronized with neighboring countries



Full correlation matrix for new cases per million in Europe



Full correlation matrix for new deaths per million in Europe

	EntertainmentVenues (new cases per 100k) Weeks												
	٨	35	ົນ	>	0	^	2	ზ	A	6	6		
<15yr	58.0579	80.745	101.673	127.461	105.945	102.982	97.6256	100.767	100.71	71.7512	63.2098		
15-24yr	205.642	256.165	315.973	374.17	300.085	285.959	246.867	232.974	207.105	158.322	141.215		
25-49yr	152.947	190.48	234.932	301.6	274.142	283.179	268.226	265.026	240.102	190.59	167.856		
50-64yr	122.642	154.49	192.018	249.57	234.785	256.421	254.44	263.69	246.132	199.246	174.72		
65-79yr	72.7789	99.195	124.873	160.617	152.236	167.805	170.565	182.924	172.934	142.229	125.215		
80+yr	103.205	136.355	173.782	235.417	245.97	264.792	256.04	306.774	327.668	315,324	305.61		

Composite mean of new cases per 100k across age groups when closing entertainment venues at week 0.

	ClosSec (new cases per 100k) Weeks										
	٨	35	າ	>	0	^	2	ზ	۵	6	6
<15yr	65.88	83.8882	98.405	103.317	81.9242	87.6757	74.0395	79.6205	76.1816	62.5179	54:0051
15-24yr	236.213	280.571	299.26	293.317	242.582	268.632	227.234	207.518	183.716	158.508	144.61
25-49уг	181.287	215.324	251.83	283.143	248.564	290.065	289.313	293.49	266.095	229.062	211.569
50-64yr	152.373	177.541	209.37	243.091	221.961	268.641	279.234	301.818	283.284	244.085	225.354
65-79yr	91.2667	108.071	130.275	154.791	140.485	171.046	184.153	206.264	193.032	168.715	157.141
80+yr	127.033	150.4	187.285	233.274	206.782	242.538	277.497	345.772	366.113	369.215	371.546

Composite mean of new cases per 100k across age groups when closing secondary schools at week 0.

	ClosHigh (new cases per 100k) Weeks											
	A	35	n	N	0	^	2	ზ	b	6	6	
<15yr	105.59	118.6	110.471	111.513	80.0538	84.69	64.3677	68.1656	61.33	47.2	37.0839	
15-24yr	302.67	316.018	272.993	262.9	214.15	252.56	200.681	178.694	156.463	130.603	108.929	
25-49уг	243.35	255.4	227.529	235.106	200.4	259.913	255.371	253.459	226.493	189.471	161.252	
50-64yr	216.96	230.273	208.85	216.812	183.908	239.083	241.923	259.319	240.17	200.229	169.929	
65-79yr	128.94	140.973	131.507	137.975	115.681	152.703	161.1	181.244	167.617	141.977	123.381	
80+yr	188.21	207.2	192.714	208.119	169.358	211.643	241.261	315.541	352.457	361.648	351.045	

Composite mean of new cases per 100k across age groups when closing higher educational institutions at week 0.

	GymsSportsCentres (new cases per 100k) Weeks											
	A	35	ົນ	>	0	^	2	ტ	A	6	6	
<15yr	69.5	98.2154	113.225	142.188	110.659	112.244	111.418	118.026	119.455	77.45	63.9156	
15-24yr	224.908	302.277	332.481	402.629	319.211	311.597	277.033	258.524	238.218	171.916	145.878	
25-49yr	174.608	229.562	257.338	344.253	303.607	317.628	316.312	320.947	304.376	223.219	191.175	
50-64yr	139.431	185.246	213.631	287.653	260.152	286.672	302.03	322.094	312.482	238.241	204.944	
65-79yr	90.1615	123.877	139.594	184.635	165.822	182.594	197.755	216.206	207.73	159.725	137.834	
80+yr	138.062	178.662	200.05	280.147	262.244	281.916	307.824	369.174	389.355	358.797	335.266	

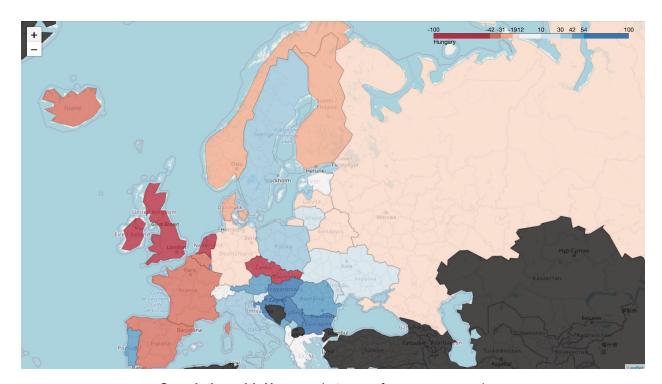
Composite mean of new cases per 100k across age groups when closing gyms sports centers at week 0.

	MassGather50 (new cases per 100k) Weeks											
	A	35	n	>	0	^	2	ზ	٨	6	6	
<15yr	44.0739	54.968	67.7259	82.8379	92.4094	99.7111	106.31	112.383	106.183	91.8	79.9634	
15-24yr	139.496	171.416	221.811	277.969	309.922	312.678	295.92	282.022	250.588	216.195	196.027	
25-49уг	104.1	125.964	165.9	233.748	295.394	315.747	315.798	326.093	302.944	260.883	233.571	
50-64yr	84.7696	102.712	133.93	191.31	252.534	286.597	298.887	315.978	301.1	267.61	238.72	
65-79yr	51.1739	62.004	79.6296	114.455	156.5	185.169	196.32	207.595	202.244	181.771	165	
80+yr	72.787	88.404	112.941	167.117	233.209	277.075	310.345	356.661	392.066	386.056	360.078	

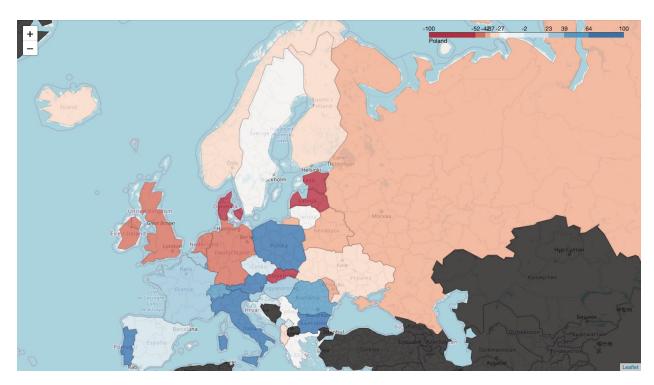
Composite mean of new cases per 100k across age groups when banning gathering over 50 people at week 0.

	StayHomeOrder (new cases per 100k) Weeks										
	٨	35	ว	>	0	^	2	ზ	٨	6	6
<15yr	69.46	77.1833	69.9875	81.9444	68.1063	85.1833	93.0889	83.3471	65.4833	57.1	42.1235
15-24yr	196.16	225.333	229.113	293.567	223.188	235.378	227.422	199.353	152.128	120.644	87.1882
25-49уг	125.06	144.567	168.487	240.089	198.469	240.089	269.606	248.382	184.628	143.061	101.606
50-64yr	96.1	119.033	142.175	208.511	180.137	230.211	278.744	270.629	203.522	153.022	107.988
65-79yr	56.76	71.7167	85.4	126.967	109.956	148.261	194.539	192.771	152.522	120.517	84.6706
80+yr	76.24	96.7333	124.025	174.467	155.238	224.606	321.689	369.324	376.617	371.628	267.682

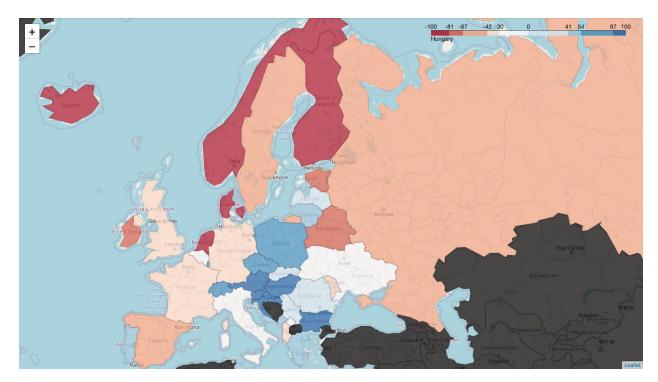
Composite mean of new cases per 100k across age groups when ordering staying at home at week 0.



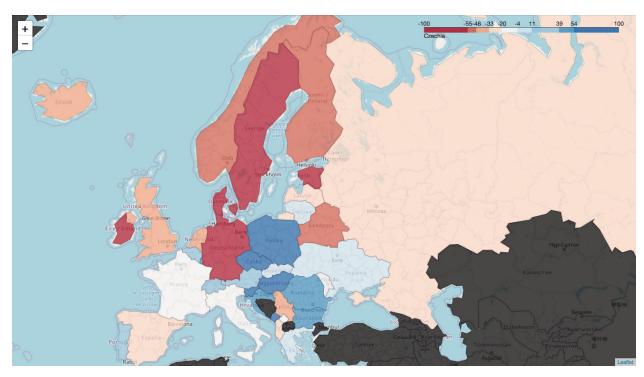
Correlation with Hungary in terms of new cases per day.



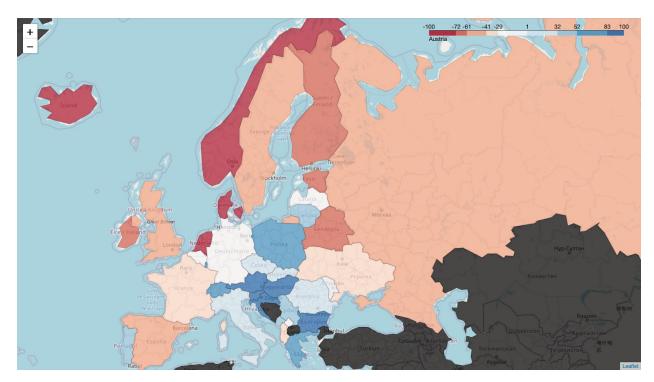
Correlation with Poland in terms of new cases per day.



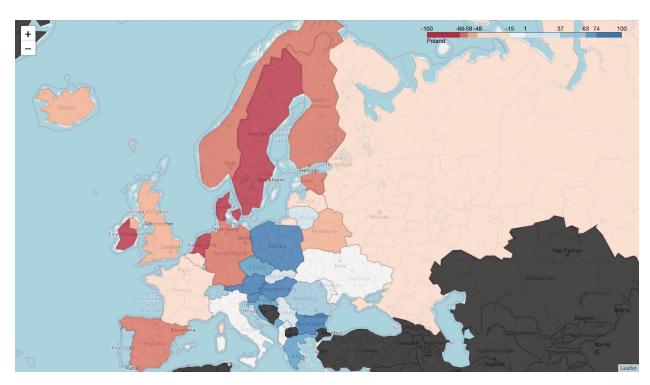
Correlation with Hungary in terms of new deaths per day.



Correlation with Czechia in terms of new deaths per day.



Correlation with Austria in terms of new deaths per day.



Correlation with Poland in terms of new deaths per day.

References

- Page, Jeremy; Hinshaw, Drew; McKay, Betsy (26 February 2021). "In Hunt for Covid-19
 Origin, Patient Zero Points to Second Wuhan Market The man with the first confirmed
 infection of the new coronavirus told the WHO team that his parents had shopped there".
 The Wall Street Journal.
- 2. "Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19)". U.S. Centers for Disease Control and Prevention (CDC). 6 April 2020. Archived from the original on 2 March 2020.
- Hutchins HJ, Wolff B, Leeb R, et al. COVID-19 Mitigation Behaviors by Age Group —
 United States, April–June 2020. MMWR Morb Mortal Wkly Rep 2020;69:1584–1590.
 DOI: http://dx.doi.org/10.15585/mmwr.mm6943e4external icon