

COVID-19 News Classification

Authors: Tegveer Ghura, Ding Ma, Gordon Ng, Yining Wang

Introduction

We aim to comprehensively analyze the correlation between the impression of the US public on the COVID-19 pandemic from April 2020 to August 2020 , as captured by a GALLUP poll, and the proportion of news report topics in the United States, China, and South Korea. By utilizing their correlation coefficients and Granger causality results, we highlight how international (and local) media coverage and public perception interact during the pandemic in contemporary society.

Methods

Data preparation:

- Dimension reduction:** Transform the impression of the public on pandemic from a 5-dimensional space into a 1-dimensional space.
- Z-score normalization:** A strategy of normalizing data to meet the requirement of using Granger Causality and avoid the influence of dimension, variable variation, and numerical value.

Correlation analysis:

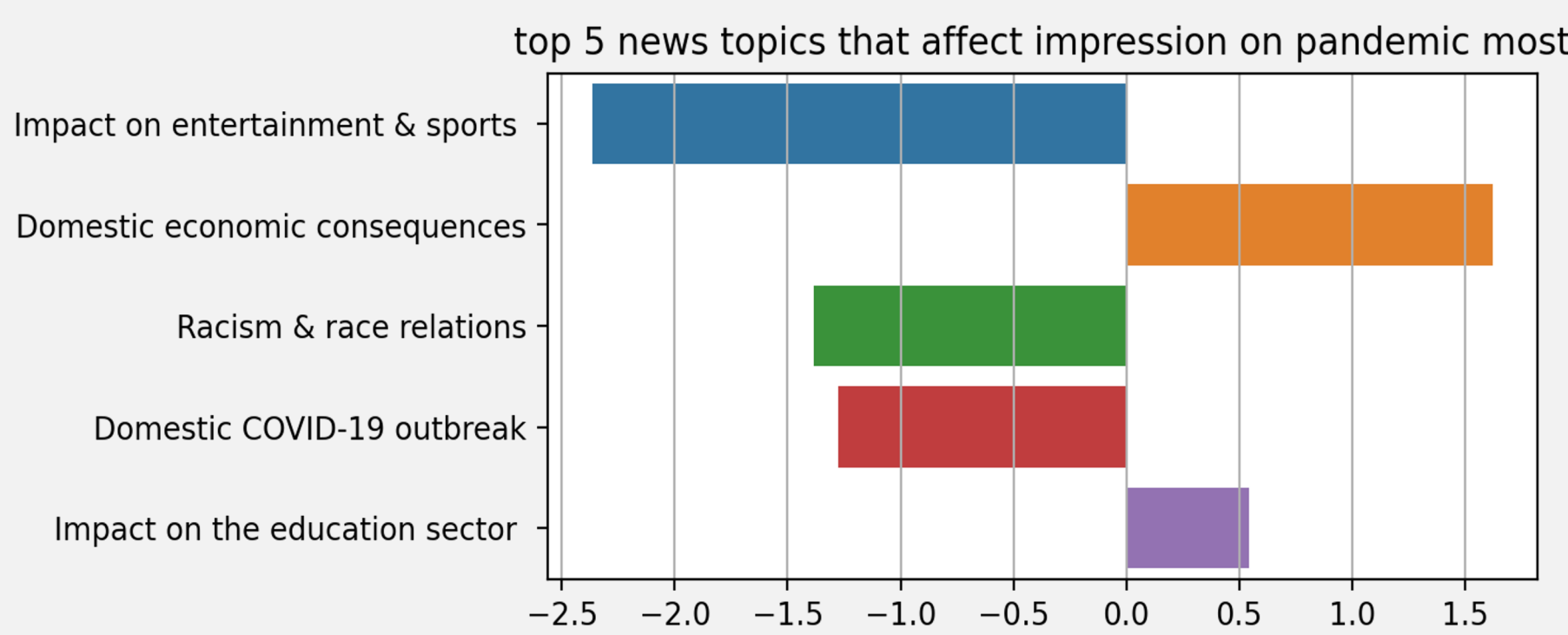
- Spearman correlation:** A non-parametric test to measure the degree of association between two variables.
- Granger Causality:** A statistical hypothesis test for determining whether one-time series is helpful for forecasting another. The F-test indicates the degree of impacts, and the P-value indicates reliability, a p-value <0.05 indicates significance.

Results

US news topic

Correlation result:

According to the result of the Spearman correlation coefficient, the most correlated with impression on pandemic news topic is the impact on entertainment & sports (negative), domestic economic consequences (positive), racism & race relations (negative), domestic COVID-19 outbreak (negative).



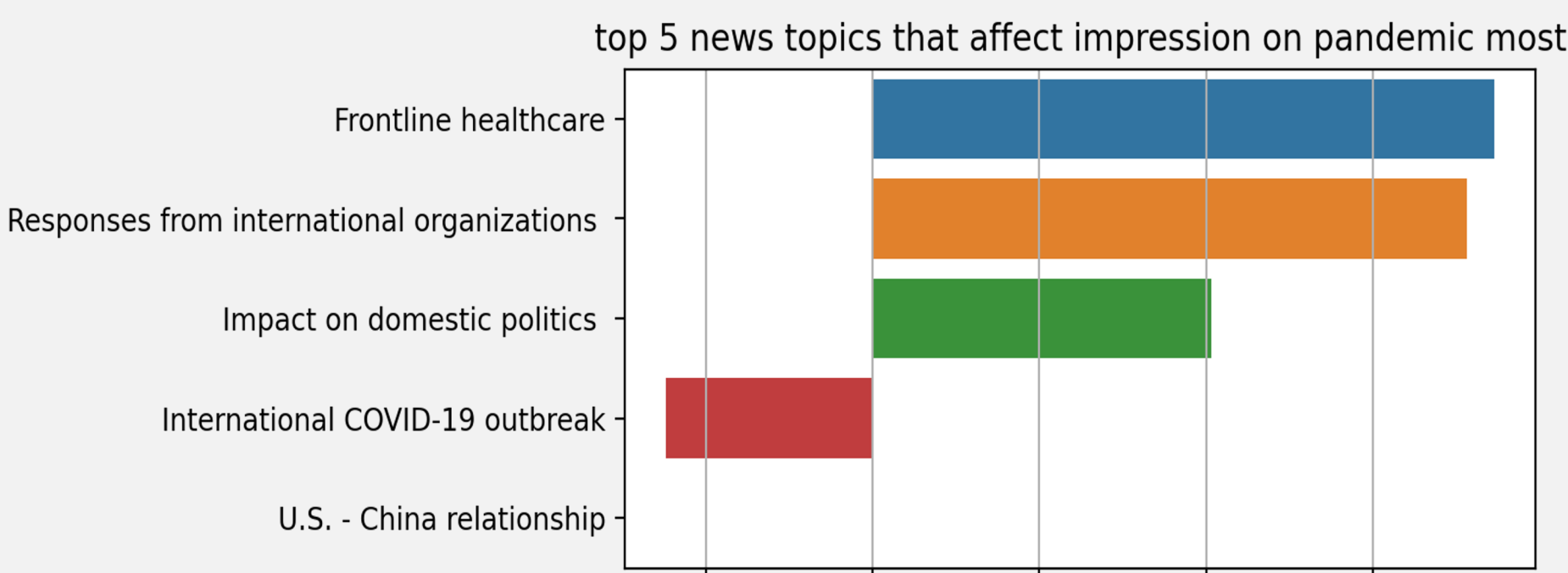
Granger causality result:

1: impact on entertainment & sports. 2: Domestic economic consequences.
3: Racism & race relation. 4: Domestic COVID-19 outbreak

	1-Lag		2-Lag		3-Lag		4-Lag	
	F-test	P value	F-test	P value	F-test	P value	F-test	P value
1	12.968	0.0029	2.3006	0.1463	2.0010	0.1925	1.4428	0.3435
2	0.7356	0.4055	4.8710	0.0305	2.4654	0.1368	4.1375	0.0757
3	3.7116	0.0746	3.8085	0.0454	3.3322	0.0770	1.7686	0.2720
4	0.5213	0.4822	1.2166	0.3332	0.5938	0.6365	6.8458	0.0292

China news topic

Correlation result:



According to the result of the Spearman correlation coefficient, the most correlated with impression on pandemic news topic are Frontline healthcare. (positive), Responses from international organizations.(positive), International COVID-19 outbreak. (negative).

Granger causality result:

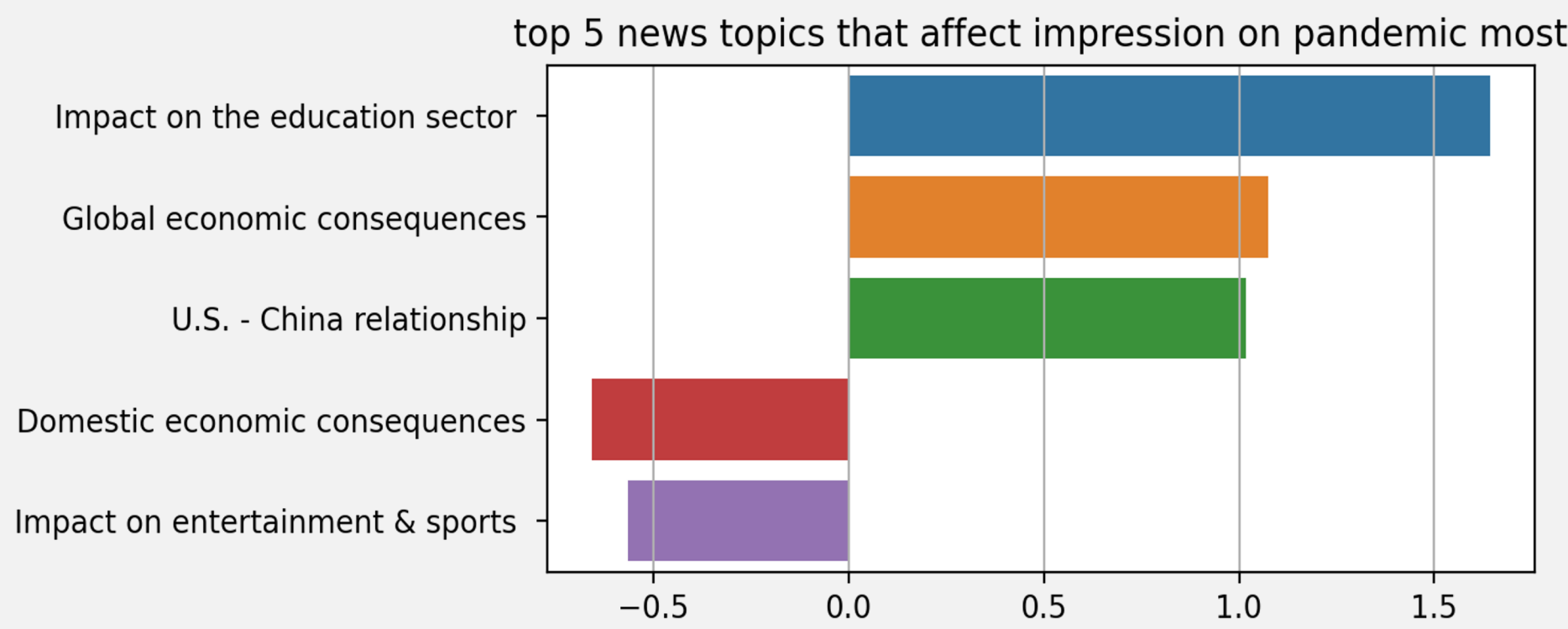
1: Frontline healthcare. 2: Responses from international organizations.3: Impact on domestic politics. 4: International COVID-19 outbreak.

	1-Lag		2-Lag		3-Lag		4-Lag	
	F-test	P value	F-test	P value	F-test	P value	F-test	P value
1	3.1907	0.0957	2.3971	0.1367	1.1116	0.3996	1.0926	0.4504
2	1.9759	0.1816	0.9279	0.4242	1.6213	0.2596	1.7378	0.2778
3	1.1861	0.2944	0.8575	0.4507	0.5050	0.6894	1.0316	0.4732
4	0.0024	0.9609	0.5196	0.6086	0.1706	0.9132	1.3223	0.3761

Korea news topic

Correlation result:

According to the result of the Spearman correlation coefficient, the most correlated with impression on pandemic news topic are impact on education sector (positive), global economic consequences (positive), US-China relationship (positive), domestic economic consequences (negative), Impact on entertainment & sports (negative).



Granger causality result:

1: impact on the education sector. 2: global economic consequences. 3: US-China relationship. 4. domestic economic consequences 5. Impact on entertainment & sports

	1-Lag		2-Lag		3-Lag		4-Lag	
	F-test	P value	F-test	P value	F-test	P value	F-test	P value
1	5.1942	0.0388	1.0595	0.3794	0.5145	0.6836	0.8185	0.5647
2	1.2215	0.2876	2.1270	0.1655	2.0227	0.1893	2.7296	0.1501
3	1.8819	0.1916	1.7130	0.2250	2.4961	0.1338	2.3165	0.1908
4	1.2067	0.2905	0.8253	0.4634	0.5055	0.6891	1.1583	0.4274
5	5.8402	0.0298	3.1422	0.0832	1.2353	0.3588	4.5482	0.0638

Granger Causality Result Analysis

US: We observed a 1-lag impact on entertainment and sports, showing that they Granger-cause US citizens' impression of the pandemic. The Spearman correlation coefficients between these factors are negative, indicating that an increase in coverage of entertainment and sports impact is followed by a decline in citizens' impression of the pandemic. Similarly, we observed a negative effect of racism, which is correlated with the "Black Lives Matter" movement that started at the end of May. Additionally, we found that the Domestic Covid-19 outbreak has negative impacts, but with a longer time lag, indicating that the US public's attitudes didn't change until the public realized the gravity of the situation. Furthermore, Trump's economic stimulation has a positive impact on people's attitudes, as reflected by the Granger results of Domestic economic consequences.

China: Although the result of Spearman correlation coefficients signifies that US public sentiment and Chinese news topics correlate with each other, the Granger Causality results disagree. Most of the news topics related to the impression of the American pandemic are about international pandemic or health care. However, there is no time-bearing effect as seen through insignificant F-values and p-values.

Korea: The Granger Causality results show that most news topics do not Granger-cause US citizens' impression of the pandemic. However, there is a 1-lag impact on the education sector and impact on entertainment & sports that Granger cause the pandemic. This result likely indicates the inconsistency of the peak of the epidemic situation between South Korea and the United States. Before the epidemic situation in the United States expanded, the epidemic situation in South Korea reached its peak, with the most negative reports on the epidemic situation. From July to August, the epidemic situation in South Korea retreated. After the outbreak in the United States became severe, the negative reports on the epidemic situation in South Korea decreased. At the same time, the impression of pandemics in the United States also reduced. Hence, they may have high Spearman correlation coefficients. However, there is no precise sequence in this relationship.

Conclusion

Our results reflect the pandemic environment from April to August 2020, demonstrating the impact of both local and international news on the American people's attitude towards Covid-19. As anticipated, most of our findings show how local news about economics, pandemics, and racism influences public attitudes. However, some of the results were unexpected, such as the influence of Korean news on education and entertainment.

To further analyze these unexpected results, we need to rely on additional data sources. Overall, our findings indicate a close correlation between news reports and the public's attitude towards specific events during the pandemic.

References

- Impression_on_coronavirus.xlsx
- Themes by Weeks.xlsx
- “Framing and Agenda-setting in Russian News: a Computational Analysis of Intricate Political Strategies”