Modeling Covid-19 in UK

Andrew Hong, Kezhen Zhao, Lily Qiao, Yuhao Chen
Supervisors: Rebecca Nugent, Peter Freeman
12/16/2020
Project Objective

Covid 19 has been disrupting human health as well healthcare all over the world, and we can expect it to continue affecting us next year.

Objectives:

- Find potential leading indicators
- Make inferences on UK Covid data
- How to predict hotspots
UK Daily Confirmed Cases

* Source: ECDC. Consistent with OurWorld Data.
Leading Indicators - Mobility
Leading Indicators - Mobility

Report mobility patterns, i.e. movement trends (% change from pre-covid baseline) over time by geography, across different categories of places such as:

- retail and recreation,
- groceries and pharmacies,
- parks,
- transit stations,
- workplaces, and
- residential.

Using MA7 (Moving Averages) to account for any recording delays over the week.

*Mobility data from Google Cloud BigQuery. Covid data from ECDC.*
Optimal Lag: retail_and_recreation = 100
Spearman Cross-Correlation: 0.8959089

Approx. 100 days
Correlation between mobility and daily cases

*Mobility data from Google Cloud BigQuery. Covid data from ECDC.*
Leading Indicators - Gov. Policy Index
# Leading Indicators - Gov. Policy Index

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Type</th>
<th>Targeted/General?</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>School closing</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C2</td>
<td>Workplace closing</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C3</td>
<td>Cancel public events</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C4</td>
<td>Restrictions on gathering size</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C5</td>
<td>Close public transport</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C6</td>
<td>Stay at home requirements</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C7</td>
<td>Restrictions on internal movement</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>C8</td>
<td>Restrictions on international travel</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>E1</td>
<td>Income support</td>
<td>Ordinal</td>
<td>Sectoral</td>
</tr>
<tr>
<td>E2</td>
<td>Debt/contract relief for households</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>E3</td>
<td>Fiscal measures</td>
<td>Numeric</td>
<td>No</td>
</tr>
<tr>
<td>E4</td>
<td>Giving international support</td>
<td>Numeric</td>
<td>No</td>
</tr>
<tr>
<td>H1</td>
<td>Public information campaign</td>
<td>Ordinal</td>
<td>Geographic</td>
</tr>
<tr>
<td>H2</td>
<td>Testing policy</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>H3</td>
<td>Contact tracing</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>H4</td>
<td>Emergency investment in healthcare</td>
<td>Numeric</td>
<td>No</td>
</tr>
<tr>
<td>H5</td>
<td>Investment in Covid-19 vaccines</td>
<td>Numeric</td>
<td>No</td>
</tr>
<tr>
<td>H6</td>
<td>Facial coverings</td>
<td>Ordinal</td>
<td>Yes</td>
</tr>
<tr>
<td>M1</td>
<td>Other responses</td>
<td>Text</td>
<td>No</td>
</tr>
</tbody>
</table>

## Index Names

<table>
<thead>
<tr>
<th>Index Name</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C7</th>
<th>C8</th>
<th>E1</th>
<th>E2</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government Response Index</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Containment and Health Index</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stringency Index</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Economic Support Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Oxford COVID-19 Government Response Tracker (OxCGRT)
Potential Lag effect:
Overall decreasing trend in indexes (June - Oct.) ---> increase in daily cases (Aug. - Nov.)
Overall increasing trend in indexes (Oct. - Nov.) ---> decrease in daily cases (Nov - now)

Stringency = C1~C8 + H1
Containment = Stringency + H2+H3+H6
Observation: cases after 87 days are corresponding better to the changes (reversely) in stringency index now.

0-lag correlation: 0.45

spearman corr: 0.613081328509765
H2: testing policy
H3: contact tracing
H6: facial coverings

0-lag Correlation: -0.18

best lag: 107
window: 7
spearman corr: -0.8727348723480409
Combined Leading Indicators

<table>
<thead>
<tr>
<th>Gov. policy indicator</th>
<th>Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workplace closing</td>
<td>workplace % change</td>
</tr>
<tr>
<td>Transit closing</td>
<td>transit stations % change</td>
</tr>
<tr>
<td>Stay at home requirement</td>
<td>residential % change</td>
</tr>
</tbody>
</table>
Cases vs. Workplace mobility/index with 7-day MA (3/16-11/30)

Correlation between workplace_closing and workplace % change: -0.6731738805662277
Correlation between workplace_closing and workplace % change (MA=7): -0.7357487707371038
This is due to transit_closing index (C5) remains unchanged during the period.

Correlation between transit_closing and transit stations % change: nan

Correlation between transit_closing and transit stations % change (MA=7): nan
Correlation between stay_home_requirement and residential % change: 0.5994838285111471
Correlation between stay_home_requirement and residential % change (MA=7): 0.7319404429433023
ARIMA
ARIMA model

What is it?

- We can use trends in the past to predict Covid cases in the future

ARIMA = Auto Regressive Integrated Moving Average

- **Auto Regressive** - Fitting a regression based on several previous timesteps
- **Integrated** - Accounts for data not being stationary (constant mean and variance)
- **Moving Average** - Forecasting based on errors from previous time step predictions
Train/test

Train: 75%: 1/23 - 9/12
Test: 25%: 9/13 - 11/30

P (auto regressive) = 6
D (integrated) = 1
Q (MA window) = 6
Model Results

Arima Test

ARIMA Model Test Error
SIR and More
SIR Model

- Modelling vs Data Mining
- JHU Novel Covid Data: include confirm, death, recover data in country level
  - Fit parameters beta and gamma to minimize RMSE
  - Poor fit due to no recover data for UK after April 13th
  - Model Assumptions not hold: close country
- Next step:
  - Adjust the model
  - Data mining approach: Bi-LSTM

Future Improvements and Acknowledgements:

➢ Leading indicators: Mobility and Government Indexes (using MA7) are highly correlated with daily cases after according for a lag.
➢ ARIMA: Perform well in fitting the data.
➢ SIR: More complete data, and model adjustments are needed to improve performance.
➢ Other Models: Currently doing research, may try it in the future.

We cannot express enough thanks to the Optum Team and CMU faculties for their continued support and encouragement regarding this project: Danita Kiser, Paul Nielsen and Professor Rebecca. The completion of this project could not have been accomplished without the support of them.
Thank You!

Any Questions?