Impact of Domestic Air Travel on Early COVID-19 Spread

Network Effects and PageRank Analysis

Sara Nicholas, Lindsey Blanks

6.268 Final Project
Massachusetts Institute of Technology

May 20, 2021

Background

- Hope to explain geographic heterogeneity in the spread of COVID-19
- Results from "The Effect of International Travel on the Spread of COVID-19 in the U.S." suggest that early ban on travel from Europe could have greatly reduced spread of COVID-19 in the US
- We focus on international flights from Italy because the early travel ban from China (Jan 31 2020) leads to negligible effects in COVID-19 spread
- We observe COVID-19 case count on the state level, on March 17, 2020 and March 31, 2020
- We construct models that include states that received direct flights from Italy and models with those states + their neighbors (connected through direct flights)
- We construct models with passengers from Italy and models with passengers from Italy + any country on Italy's border
- We ran both log-linear and Poisson regressions. We found similar trends so we will show the results from the log-linear regressions only in this presentation for clarity

Baseline Model

c_i: COVID-19 cases in state i
x_i: incoming passengers from Italy to state i (normalized)
t_i: tests conducted in state i (normalized)

$$\log(\mathrm{c}) = \alpha \mathrm{x} + \eta \mathrm{t} + k$$

y_i: incoming passengers from Italy and bordering countries (France, Switzerland, Austria, Slovenia) to state i (normalized)

$$\log(\mathrm{c}) = \gamma \mathrm{y} + \eta \mathrm{t} + k$$

Baseline Model - Results

$$\log(c) = \alpha x + \eta t + k$$

$$\log(c) = \gamma y + \eta t + k$$

Italy

March 17:
$\alpha = 3.57$
$\eta = 4.14$
r squared: 0.458

March 31:
$\alpha = -1.31$
$\eta = 7.84$
r squared: 0.561

Italy + Neighboring Countries

March 17:
$\gamma = 4.08$
$\eta = 3.60$
r squared: 0.534

March 31:
$\gamma = 1.72$
$\eta = 5.36$
r squared: 0.568

Neighbors Model

c_i: COVID-19 cases in state i
x_i: incoming passengers from Italy to state i (normalized)
t_i: tests conducted in state i (normalized)
w_{ij}: passengers from state i to state j

$$m_i = \sum_j w_{ji} x_j \qquad n_i = \frac{m_i}{\max\{m\}}$$

$$\log(\mathrm{c}) = \alpha \mathrm{x} + \eta \mathrm{t} + \sigma \mathrm{n} + k$$

Neighbors Model - Results

$$\log(\mathsf{c}) = \alpha \mathsf{x} + \eta \mathsf{t} + \sigma \mathsf{n} + k$$

March 17:
$\alpha = 2.85$
$\eta = 3.67$
$\sigma = 2.44$
r squared: 0.545

March 31:
$\alpha = -1.01$
$\eta = 6.70$
$\sigma = 2.13$
r squared: 0.614

PageRank Model

c_i: COVID-19 cases in state i
x_i: incoming passengers from Italy to state i (normalized)
t_i: tests conducted in state i (normalized)
w_{ij}: passengers from state i to state j

$$r_i = \sum_{j \neq i} \frac{w_{ji}}{\delta_j} r_j + \beta_i \qquad p_i = \frac{r_i}{\max\{r\}}$$

(1) $\beta_i = 1 \ \ \forall \beta$
(2) $\beta_i = x_i$

$$\log(\mathrm{c}) = \alpha \mathrm{x} + \eta \mathrm{t} + \mu \mathrm{p} + k$$

PageRank Model - Results

$$\log(c) = \alpha x + \eta t + \mu p + k$$

$\beta_i = 1$

March 17:
$\alpha = 2.76$
$\eta = 3.14$
$\mu = 2.33$
r squared: 0.565

March 31:
$\alpha = -0.80$
$\eta = 6.36$
$\mu = 2.03$
r squared: 0.630

$\beta_i = x_i$

March 17:
$\alpha = 1.94$
$\eta = 3.17$
$\mu = 2.37$
r squared: 0.571

March 31:
$\alpha = -1.47$
$\eta = 6.19$
$\mu = 2.09$
r squared: 0.635

Conclusions

- Considering network effects improves the accuracy of the models.
- On March 17, travel from Italy has high significance. By March 31, domestic travel and testing become much more important.
- Including other parameters (population density, demographics, etc) could further tune the model.
- Heterogeneity within a state creates noise, looking at county or city level data could eliminate some of this noise.
- Domestic air travel played large role in COVID-19 spread in mid-late March 2020. Earlier implementation of social distancing measures and mask mandates on flights, as well as domestic travel restrictions, may have helped curb this early spread through the U.S.

Questions