COVID-19 Government Response Framework Affect Weekly Outbreaks in Ontario 2020

Zhen Xia

December 21st 2020

Abstract

COVID-19 conducted negative impacts of community-based public health hence it is necessary to investigate what explore what policies adopted by the Ontario government and whether the policies are effective. This research utilizes data from four main regions of Ontario, which are Toronto, Ottawa, York and Peel, to explore detailed weekly new outbreak data and the state of the region at that time. The demographic characteristics of refreshed weekly data are related to the government's restrictions on the public. The result of ths research shows that after a fixed time line May 17th and the governments unblock the cities and districts, and this will result in an increase in new outbreaks.

Keywords

COVID-19 New Outbreaks Ontario Government Policy Response Framework Unlock City

Introduction

Statistics analysis is a useful and appropriate method in public health researches, especially during the COVID-19 epidemic that the world had been experienced in the past several months. The transmissibility and harmfulness of COVID-19 require almost every large institution to refresh the relevant statistics as soon as possible to ensure that the public and the governments are able to be in preparation to the outbreak. In Ontario, where the University of Toronto located, it is reasonable to assume causal links existing between the number of new outbreaks and other various variables, such as area, population size, population density, government response frameworks and so on. Thus, if the laws that causal relationships follow could be figured out, which is what this research endeavor to talk about, the governments might be able to manage the outbreak more effectively to deal with the current situation.

To achieve this goal, one of the statistical techniques that can be adopted is the Difference-in-Difference. This is used in statistical analysis to quantize the effect of a treatment using observational data from two groups, which is 'control group' and 'treatment group'. The mechanism that DID works is to calculate and compare and average change over time of control group and treatment group separately. By this mean the data collected for this research is an observational data. The observational data are achievable and authentic, both are indelible advantages in comparison to experimental data.

The data comes from the websites of the four main districts of central Ontario, which includes the districts of Toronto City, Ottawa City, York district and Peel's district. Dataset would be simulated and then performed DID on it. The treatment group in this research would be the time period which the COVID-19 Response Framework is implemented in one specific region and control group would be the time period in which the

COVID-19 Response Framework is not implemented in another region. Despite this, area, population size and population density would be added in the model to eliminate their effects to the model.The result of the DID is provided in the Results section and the following talk about the causal inferences between COVID-19 increases and all variables is put in Discussion section.

Data

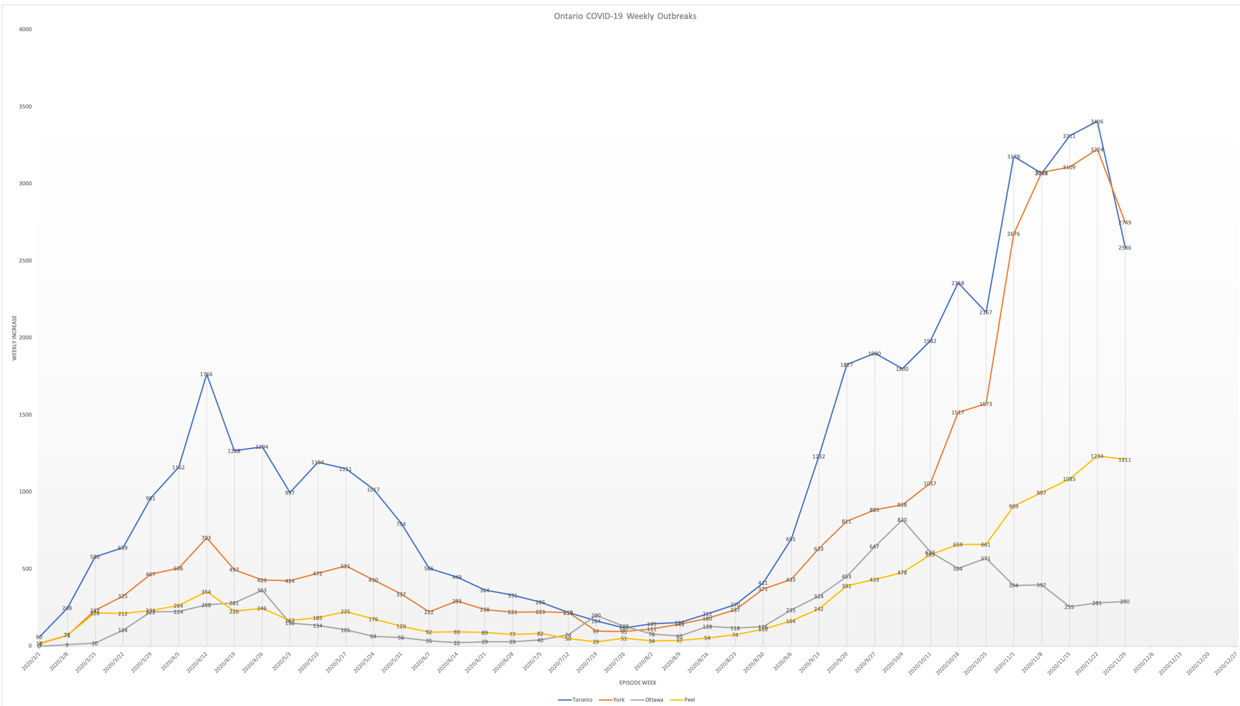

Figure 1: COVID-19 Trends

Weekly_increase is the new outbreak data. The frontline health workers are obligated to report and confirmed or suspect COVID-19 cases to PHAs meanwhile the PHAs are shall connect the frontline health workers and provincial laboratories to keep in touch throughout the whole treatment. As a consequence, non-response appears rarely except asymptomatic individuals, that is people who are infected without any obvious symptoms, can hardly be tracked. There are also a small group of people who have died in the complication caused by the corona virus and will never be found infected without judicial autopsy. A time variable starting counted weekly from March 1st is episode week. Policy variable is the response framework of regional governments for local residents. Before November, the Ontario government separated three phases the state of the province, which is phase 1: lock down, phase 2: half-closed and phase 3 unblocked. After November, Ontario issued a new framework, that is gray: lockdown, red: control, orange: restrict, yellow: protect, and green: prevent. Policy is a dummy variable has qualitative attribute as strictly lock down. The reason is only large-scale restrictions on any social activities can greatly reduce the possibility of infection from close contact for the public. All data graphs show significant increase from the beginning of March and then reach a peak in mid-April, followed by slowly fall until August. In October, it rose to peak value again and displays a trend of continuous increase.

These data were sourced from the official websites of the four major districts in Ontario, that is Toronto Public Health, Region of Peel, Ottawa public health, and COVID-19 in York Region. Toronto Public health receives the COVID-19 reports from the local public health departments sending by the laboratory. York and

Peel extracts the dataset from CCM, which is the Public Health Case and Contact Management Solution from Ontario Ministry of Health, for they are both regional municipality of Ontario. As well as Ottawa's official datasets are based on the information from the COD, that is the COVID-19 Ottawa Database reporting system updating 2 p.m. every day. The population would be all people infected with the corona virus in Canada. Among the target population, I choose the sampled population, all the new outbreak from March First to the latest update date,November 29th, from the sampling frame of all the daily increases of COVID-19 cases by now. This dataset is almost the most accurate and convincing data available to the public because of its official nature. Inevitably shortages include the existence of non-response and variation of data collection methods between regions, leading to dividing contemporaneous data into different time periods.

Model

General Model

The model section builds a difference-in-difference model to investigate the number of weekly_increase using 3 factors using data from March to December.

$$y = \beta_0 + \beta_1 x_{time} + \beta_2 x_{treatment} + \beta_3 x_{time\text{Ö}treatment} + \beta x + e$$

y is the dependent variable in this model. It refers to weekly increase data in this research.

β_0 is the interception which represents when all categorical variables are at baseline, the number of new outbreaks per week.

β_1 is a dummy variable for time.The weekly increasing data is divided by before or after a specific time line, which is May 17th, the date that Peel's district is blocked down.

β_2 is the dummy variable for treatment. The cities should be completely locked down by the government or not.

β_3 is the dummy variable for time*treatment . It represents whether the district is completely locked down after May 17th or not.

Part Model

The model section builds a logistic model to investigate the number of weekly_increase using 3 factors using data from March to August as a complete epidemic cycle.

$$y = \beta_4 + \beta_5 x_{time} + \beta_6 x_{treatment} + \beta_7 x_{time*treatment} + \beta x + e$$

y,x,e and β_4 to β_7 which have exactly the same meanings as the coefficients and variables in the general model correspondingly, are omitted.

Categorical variables:

x_{time},$x_{treatment}$and $x_{time*treatment}$ are dummy variables,

x_{time} is the dummy variable of before or after a specific timeline.

$x_{treatment}$ is the dummy variable of treatment group or not.

$x_{time*treatment}$ is the covariate variable defined by treatment group and after timeline is performed at the same time.

x is the fixed effect added into the model,including citizen's execution, the population density of the area, location, evolution of virus and so on.

e is the remaining residuals.

$$x_{time} = \begin{cases} 1 & if\ time\ is\ after\ May\ 17th \\ 0 & otherwise \end{cases}$$

$$x_{treatment} = \begin{cases} 1 & if\ the\ district\ is\ completely\ reopen \\ 0 & otherwise \end{cases}$$

$$x_{time*treatment} = \begin{cases} 1 & if\ the\ district\ is\ completely\ reopen\ after\ May\ 17th \\ 0 & otherwise \end{cases}$$

Result

Interpretation of β_k:

Table 1: Difference-in-Difference Results

Variables	Coefficients	P_value
(intercept)	226.82	0.00118 **
time	-170.99	0.19817
treatment	-104.32	0.91525
time * treatment	115.39	0.10735
general model	-	0.1311

P-value for the general model using all data is 0.1311, which is larger than 0.05, shows non-significance for the general model, hence no significant evidence to show there exists a causual inference.

β_0: When no government framework policy applied to the cities, weekly outbreaks increases before the speicific time line, May 17th.

β_1: When the data collecting area is completely reopen, weekly new outbreaks decreases at a higher rate.

β_2: When the data collecting area is after May 17th, weekly new outbreaks decreases at a higher rate.

β_3: When the data collecting area is completely locked down and after May 17th, weekly new outbreaks increases at a slower rate.

Table 2: Difference-in-Difference Results

Variables	Coefficients	P_value
(intercept)	226.82	$<$ 2e-16 ***
time	-170.99	2.45e-07 ***
treatment	-12.70	0.00265 **
time * treatment	275.27	0.00975 **
part model	-	8.558e-10

The part model built using data from March to August has the p-value for 8.558e-10, it shows significance and is significant much smaller than that of general model. Therefore, futrher analysis will focus on the part model and corresponding data.

β_4: When no government framework policy applied to the cities, weekly outbreaks increases before the speicific time line, May 17th until August.

β_5: When the data collecting area is completely locked down, weekly new outbreaks decreases at a higher rate before August.

β_6: When the data collecting area is after May 17th, weekly new outbreaks decreases at a higher rate before August

β_7: When the data collecting area is completely locked down and after May 17th, weekly new outbreaks increases at a slower rate before August.

Discussion

The data before August was selected for the part model DID analysis by reason of the epidemic from March to August is a complete cycle of COVID-19 spreading to lead to a large-scale infection and then gradually subsiding. A DID analysis in one accomplished cycle should be compelling. The failure from the former general model also proved this point to a certain extent. Due to the experience of the previous corona virus epidemic, the government might respond to the new outbreaks differently in the second half of the year. Indeed, a new treatment policy called response framework is applied during November, when the second wave of the epidemic is coming, that is to set different color standards for each region, and apply standards to the distinct region according to the specific conditions. Factors like this can cause other fixed effects in the model to be superimposed, making general model not showing unreliable.

In this research, I used the data from Ottawa and Peel, and made a timeline on May 17th to satisfy the parallel trend assumption which is externally important to confirm validity of DID analysis. Before May 15 as the dividing line, the increase rates of weekly new outbreaks in these two regions before are almost the same. And after that date, Ottawa was lifted from the state of emergency when Peel District and Toronto City still remain blocked. The difference between treatment groups and control groups is constant over several weeks before May 17th. Data from other regions do not satisfy this assumption, so they are not included in the DID model preventing bringing wrong bias of estimating causal effects.

Although the government has been using various levels of control measures to adjust and tighten public health measures. Starting from middle of March, large districts in Ontario are declared as a state of emergency and for simplicity, it is called Stage 1 as the epidemic declined, the city was gradually opened, which is stage 2. After the cycle is over and the cities are fully opened, it will be called stage 3. After November, Ontario government introduces Ontario's COVID-19 color-coded system. From restrictions allow activities in Stage 3 without vaccine or treatment to massive restrictions and declaration of emergency, government measures are divided into five levels. This research uses a dummy variable that directly sets the stage1 status and the grey color response framework as a factor of "block down state".

The considerations, of course, include too many levels is difficult to categorically divide the intensity of government policies. Not to mention, most people will not strictly abide by the restrictions set by the government. Only when a severe epidemic has endangered everyone's life as well as the government has to issue the highest-level restraint order will that be effective enough. Reports about the general public ignore the restraint and raise the risk of infection could be seen everywhere in the world.

It will also affect the results of some DID analysis, for the binding force of the government response policy is weak to some extent. Time shows a negative coefficient in the part DID model, which shows the new cases after May 17th are increasing in a slower rate. After the corona virus spreads out, the policies performed by the government and the self-protection adopted by the general public have all worked to stop virus transmission. Moreover, the total population in any area always have an upper limit. Consequently, it is inevitable that the new outbreaks would eliminate as time goes by. The mainstream society hold an opinion that high temperature and humidity may reduce the activity of the virus, thereby slowing the spread of the coronavirus. It is true that our data chart conforms to this conclusion. The resource of this opinion is the last large-scale epidemic caused by homologous SARS virus in 2002 and an article published in Science magazine by the Princeton Environmental Institute this year believes that the current version of virus does not show trends to be affected by summer season. A more reliable statement from researchers at Mount Auburn Hospital in Cambridge reflects that seasons do affects the COVID-19 cases but not too much.

What is unusual is that the treatment variable also reveals a small negative coefficient. One of the reasons for this phenomenon might be that the treatment does not take effect immediately. When the new outbreaks have dropped to tens, source of infection would reduce and thus reopening the city would not lead to another period of infection. Besides, the population densities of Ottawa and Peel are 334.81 and 1109.35 per square kilometer respectively. Such small population density causes the effect of closing and unblocking policies to be less powerful. In addition, neither of these two cities have international airports like Toronto, which is shown in the picture, has always been the area with the highest number of new outbreaks all the year. Last but not least, the latest news appeals many countries have recently discovered the spike mutation D614G. It is difficult to determine the transmission mode of D614G and whether the current policies taken by the governments would continue to be efficient.

The coefficient of time-treatment interaction DID model shows that the coronavirus has progressively got back after the government reopen the district. This is inevitable apparently for the corona has not been eliminated worldwide, especially in United States bordering Canada. The inattention of the Ontario government and people's exclusion of "cumbersome" jobs like wearing masks would assure the virus make a comeback. This cannot fully explain the unexpected huge amount of outbreaks in December, and the influence of the spike mutation D614G in it is unknown.

Weakness

When data from each district is collected, the difference in various health and data collecting systems causes a slight difference in data counting. Variant of coronavirus D614G appears while the current research is only for COVID-19 popular in 2020. Timeliness will greatly lead to the limitations of this research. DID analysis have requirements on treatment group and control group to satisfy parallel trend assumption, which makes only a small part of the data meet this requirement. It is almost impossible to draw general conclusions on all of the data from all districts.

Next step

Extracting the data of each city from the official Ontario province data might reduce the errors caused by the different systems of each districts. The scope of the survey may be expanded to Canada or even the world in order to find a dataset that is more appropriate for the hypothesis, and at the same time, a broader conclusion can be drawn.

Reference

1. RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.
2. Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra
3. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

4. Hao Zhu (2020). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.2.1. https://CRAN.R-project.org/package=kableExtra
5. Paul-Christian Burkner (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal, 10(1), 395-411. doi:10.32614/RJ-2018-017
6. Paul-Christian Burkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01
7. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. Baker RE, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. Science. 2020 May 18.
8. JavaScript is required to view this site. (n.d.). Retrieved December 22, 2020, from https://www.ontario.ca/page/covid-19-response-framework-keeping-ontario-safe-and-open
9. Plante, J., Liu, Y., Liu, J., Xia, H., Johnson, B., Lokugamage, K., . . . Shi, P. (2020, October 26). Spike mutation D614G alters SARS-CoV-2 fitness. Retrieved December 22, 2020, from https://www.nature.com/articles/s41586-020-2895-3
10. Nielsen, K. (2020, December 17). A timeline of the novel coronavirus in Ontario. Retrieved December 22, 2020, from https://globalnews.ca/news/6859636/ontario-coronavirus-timeline/
11. City of Toronto. (2020, December 16). COVID-19: Status of Cases in Toronto. Retrieved December 22, 2020, from https://www.toronto.ca/home/covid-19/covid-19-latest-city-of-toronto-news/covid-19-status-of-cases-in-toronto/
12. Main. (n.d.). Retrieved December 22, 2020, from https://www.peelregion.ca/coronavirus/case-status/
13. Daily COVID-19 Dashboard. (n.d.). Retrieved December 22, 2020, from https://www.ottawapublichealth.ca/en/reports-research-and-statistics/daily-covid19-dashboard.aspx
14. Canada, P. (2020, September 04). Government of Canada. Retrieved December 22, 2020, from https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/health-professionals/interim-guidance-cases-contacts.html
15. Difference in differences. (2020, November 05). Retrieved December 22, 2020, from https://www.tellingstorieswithdata.com/06-03-matching_and_differences.html

Appendix

Code and data supporting this anaylysis is available at: https://github.com/ZhenXia9836/COVID-19-Response-Framework-affect-weekly-outbreaks-in-Ontario-2020