Identifying the influences of Early-Policies to the COVID-19 transmission and Deaths

Tianhao Cao

March 18, 2022

Abstract

Due to the transmission of COVID-19, the countries around the world implemented their policies to against the damages that COVID-19 may bring to them, however, the different strategies lead to varieties of results. In this essay, I estimated the early policies effects on COVID-19 transmission and damages through 4 dependent variables: Daily new cases, cumulative cases, daily deaths, cumulative deaths. The data is provided by World Health Organization and Our World in Data. I implemented two different Difference-in-Difference models to estimates the effects of Mandatory policies, Optional policies, and no-control policies with 3 groups of analysis. I found that mandatory-policy has the most significant impact on limiting and decreasing COVID-19 transmission and damages, then followed by optional-policy, and no-control policies are the least effective. I also found that the optional policies will somehow increase the cumulative deaths compare with no-control policies.

1 Introduction

Should government officials take responsibility for those infected or killed by COVID-19? COVID-19, an infectious disease, the first case was discovered in China in 2019 December. Although the Chinese government's and World Health Organization(hereinafter referred to as WHO)'s attentions were drawn, it did not cause a general concern among the world society. Soon, With the widespread of COVID-19, the administrations of different countries eventually released their early policies to limit the transmission, including mandatory or optional keeping social distance, wearing medical masks, clean hands with sanitizer. However, the results of such policies vary due to the adaptive restrictions and coercive measures, which such "adaptive policies" were made based on the dilemma between the National Economy and citizens' lives.

The essay will focus on the effects of the early COVID-19 policies released by these countries. Using WHO COVID-19 data, I will analyze the infected cases and death within six countries over time, Canada, China, Japan, India, the United Kingdom, and the United States of America. To analyze the effects, I implemented the Difference in Difference method to the data to examine the impact of policies on four measurements: new cases, cumulative cases, new deaths, and cumulative deaths. Furthermore, according to [RMRG+20] data, I am able to determine countries' policies' stringency at the early stage of transmission. Therefore, I am able to categorize these countries' policies into three classes: 1. Mandatory Countries, where the policy strictly restricted travel, enforced citizens to wear medical masks, mandatory keep social distance, etc., China is seen as a mandatory country; 2. Optional Countries, where the policy restricted travel, advised citizens to wear masks, keep social distance, etc. Canada, Japan, and the United Kingdom is seen as the optional countries; 3. No-control Countries, where the policy restricted the travels but did not provide other suggestions, the United States of America, India are seen as a no-control country. By investigating the data, I designed three groups of estimations: The first group utilizes Mandatory countries as the treatment group, other-policy countries as the control group; the Second group utilizes the mandatory countries as treatment groups, optional countries as the control group; the Third group utilizes optional countries as the treatment group, no-control country as the control group. The effects of the policies will analyze all of the above estimations through four dependent variables.

2 Literature Review

The impacts of the policies toward infectious disease are wide discussed by various scholars, from political aspects to economic aspects. Nonetheless, the results maintain huge differences based on their different starting points and goals to achieve, but in the limiting transmission aspect, scientists familiar believe that restrictions are needed to avoid further loss. Chinese scientists team[TLL+20] implemented a general linear model (adjusted by Akaike Information Criterion) estimation, which indicates that the Chinese government's early restrictions decreased at least 74,400($\pm$15,6000) possible infections in 15 days, which is approximately "96% reduction on the total number of cases expected in the absence of interventions." (Huaiyu Tian et al. 2020, Page 5), considering the lingering effects of the reinfection, the effects of the policy exceeded the estimation. Furthermore, the U.S. scientist team estimated the impact of the policy that the fifteen states plus Washington, D.C. published on April 1 about mandatory wearing masks, [LW20] indicates that "because of these mandates, 230,000–450,000 cases may have been averted by May 22" (Lyu and Wenhby 2020, page 1422). Additionally, Imad A. Moosa[Moo20] also provided an empirical statistical analysis about the social lockdown and keeping social distance. She concludes that "The results show that social distancing is effective, even though it is difficult to separate the effect of government measures from that of social distancing triggered by personal initiatives" (Moosa 2020, Page 6292) and "First, countries that have not imposed lockdown and social distancing have performed poorly... Second, countries that have imposed lockdown either late or without stringency have performed poorly. Third, countries that have imposed social distancing at lower levels of cases have done better than countries that have imposed social distancing at higher levels of cases. It seems, therefore, that social distancing does matter." (Moosa 2020, Page 6300); she pointed out that keeping social distance and social lockdowns maintain positive effects on decreasing the active cases within the country.

The estimations above are consistent with the primary goal of COVID-19 early policies, limiting the infection cases and avoiding further disease outbreaks, especially due to the lingering reinfection effects.

3 Data

3.1 WHO Data

World Health Organization(hereinafter referred to as WHO) is a non-profit global organization that intend to promote global health and respond to health emergencies. Due to its responsibility of responding to health emergencies, which collected the COVID-19 data of different countries in time series, which provides a basic data set for me to analyze. The data is divided into eight sections: Date Reported, Country Code, Country, WHO region, New cases, Cumulative cases, New deaths, and Cumulative deaths. Also, the data contains 1511 observations of 6 countries: Canada, China, India, Japan, the United Kingdom, and the

United States of America. The data is well-organized, but there is bias in the data, which is that most of the countries can not afford the medical pressures that COVID-19 bring to them hence it led to the late-record and late calculation due to short of hands, the late-record brings bias to the data set yet I can not ignore the extreme value since it is correlated with other variables, which indicates that the ignoring will result in a larger error for the regression. By observing Figure 1, we are able to have an intuitive view of the situations, it has been noticed that there is a negative increase in U.S. daily deaths, CDC [CDC22] announced that this is due to the mis-record of the actual death reasons. As we can see in the figures, the United States of America and India maintain relatively high COVID-19 cases and deaths; China, India, Japan, and the U.K. experienced relatively low pressure instead. Furthermore, to limit the influences of population, figure 2 measured the above graphs in terms of case/deaths per million people through the formula $\left(\frac{cases}{population}\right) \times 1000000$. After conducting the graph, we can see the data of the U.K. and Canada increased exponentially, this is due to the large populations in other countries enlarged the scale of the recorded variables.

3.2 Our World in Data

Our World in Data is an individual non-profit organization that intends to "Against the world's largest problems" by researching data. The organization provides a set of data that indicates the policy stringency of each country through 2020/01/02 to present; the stringency index is based on nine metrics from school closures to stay at home requirements, [HAG+21] states that the data is being spot-checked, which means that the data maintains "high degree

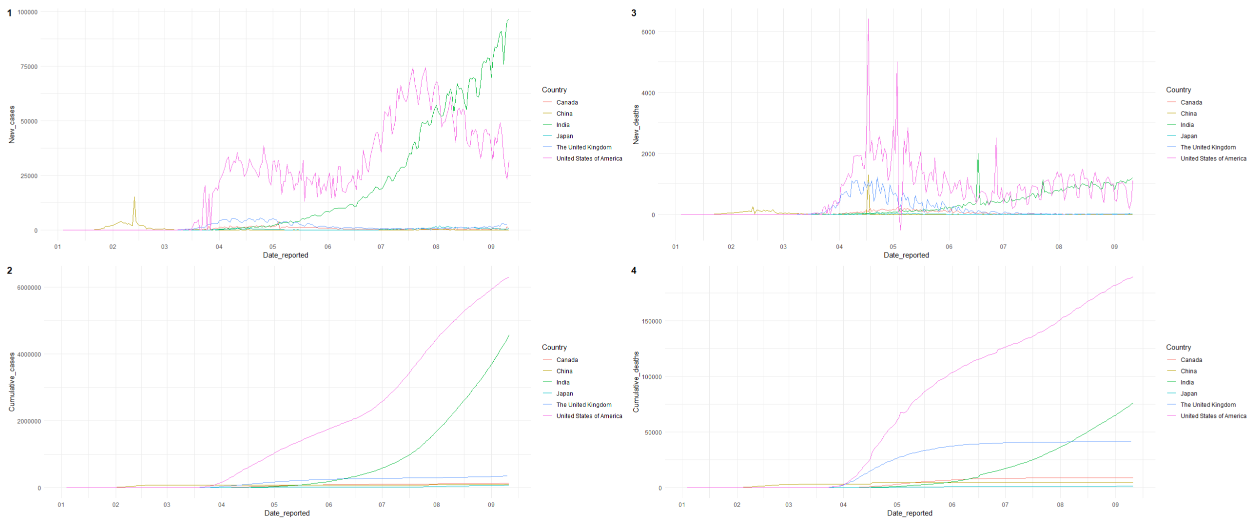

Figure 1: Line graph of cases and deaths

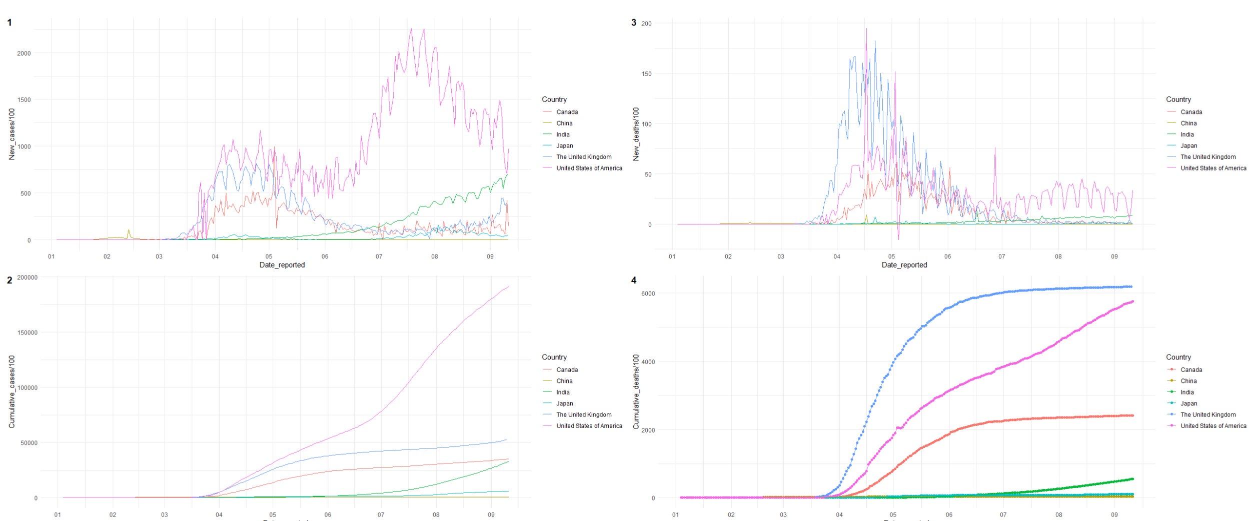

Figure 2: Line graph of cases and deaths in terms of per 1 million people

of accuracy in the initial data collection," which we don't need to consider the inaccuracy of it. Using their data, I am able to allocate the treatment group, control group, before treatment time, and after treatment time for each country that I estimate.

4 Method

I implement two Dif-in-Dif research designs to estimate early policies effects on decreasing further COVID-19 transmissions, where daily new cases, daily deaths, cumulative cases, and cumulative deaths are used as the dependent variables to estimate the effects. The analysis's motivation is that multiple literature works indicate that early policies have huge impacts on the COVID-19 situation. Meanwhile, the stringency indexes suggest that the policy contains numerous aspects. Therefore, they are considered an aggregate policy, which means that no other same-objective policies are excluded, resulting in inaccuracy to our estimation. Furthermore, since the independent variables should have no huge differences before the treatment for both treatment group and control group for implementing DID, I use per 1 million people as the unit to estimate the regression so that I can ignore the influences that population brings to the estimation.

4.1 Regular DID approach

$$y_{g1it} = \lambda_0 + \lambda_1 After_t + \lambda_2 Mandatory_{1i} + \lambda_3 (After_t * Mandatory_{1i}) + \eta_{it} \tag{1}$$

$$y_{g2it} = \rho_0 + \rho_1 After_i + \rho_2 Mandatory_{2i} + \rho_3 (After_i * Mandatory_{2i}) + \nu_{it} \tag{2}$$

$$y_{g3it} = \psi_0 + \psi_1 After_i + \psi_2 Optional_i + \psi_3 (After_i * Optional_i) + \phi_{it} \tag{3}$$

The first DID research is based on the basic DID estimation model, where I set two different groups. The first group uses Mandatory Country as the treatment group and Optional Country as the Control group, where the second group uses Optional Country as the treatment group and No-Control country as the Control group.

Equation (1) represents the estimation model for group 1, where y represents four indicators: New cases per million people, Cumulative cases per million people, New deaths per million people, and Cumulative deaths per million people. $After_i$ is a dummy variable that indicates whether the country had released its policy, $Mandatory_i$ indicates the country's policy is strict or not, eta_it is the error term. My interest focuses on λ_3, which estimates the mandatory policy's effect on COVID-19 further transmission and deaths. Equation (2) represents the estimation model for group 2, same with equation (1), the interest is focused on ρ_3, where it estimates the effects of the mandatory policy on COVID-19 further transmission and deaths, note that in equation (2), the data of the country that implements no-control policy will be excluded. Equation (3) represents the estimation model for group 3, ψ is my estimation interest, in this estimation, the data of the country that implements mandatory policy will be excluded.

4.2 Modified DID approach

$$y_{it} = \alpha_0 + \alpha_{Country}Country + \alpha_1 After_i + \alpha_2 Mandatory_i * After_i + \alpha_3 Optional_i * After_i + \mu_{it} \tag{4}$$

The second DID research is modified based on the basic DID estimation model, where I imposed Country dummy variables to estimate the effects of different policies. The benefit of using such regressions is that I am able to control the time-invariant country-specific factors, so the results will be less biased than the basic DID model.

Equation(3) represents the estimations model. $Country$ are six dummy variables that indicate each country, $After_i$ is a dummy variable that indicates whether the country had released its policy, mu_it is the error term. The estimates of interest here are α_2 and α_3, they estimates the effects of the mandatory policy and optional policy.

5 Result

5.1 Regular DID Approach

Table 1: Normal Difference in Difference model for first group

	Dependent variable:			
	‘New_cases/100‘	‘Cumulative_cases/100‘	‘New_deaths/100‘	‘Cumulative_deaths/100‘
	(1)	(2)	(3)	(4)
after	350.497*** (21.972)	28,285.040*** (1,859.727)	16.704*** (1.351)	1,895.201*** (104.553)
mandatory	−3.902 (77.714)	−174.852 (6,577.587)	−0.068 (4.777)	−17.452 (369.787)
mandatory:after	−348.148*** (82.004)	−27,720.010*** (6,940.734)	−16.572*** (5.041)	−1,867.701*** (390.203)
Constant	4.324 (18.505)	176.212 (1,566.254)	0.081 (1.138)	17.489 (88.054)
Observations	1,511	1,511	1,511	1,511
R^2	0.193	0.176	0.125	0.234
Adjusted R^2	0.191	0.175	0.123	0.233
Residual Std. Error (df = 1507)	354.024	29,964.190	21.764	1,684.565
F Statistic (df = 3; 1507)	119.822***	107.429***	71.535***	153.591***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1 indicates the estimates for running equation (1). the $\lambda 1$ variable which estimates the effects of $After$ on dependent variables, the reason that coefficients of $after_i$ turn out to be influential to the dependent variables is that COVID-19 is an infectious disease that has an incubation period, this leads to a phenomenon that most of the cases were recorded and out-breaked after the policy is implemented, in other words, the $after_i$ is overestimated. However, it does not influence the estimation since our interest of estimation is focused on λ_3, which attached $Mandatory * After$ variables, as we can see in table 1, the mandatory policy is highly correlated with the dependent variables after it was posted, same with the λ_1, the effects of the mandatory policy is also lagged due to the characteristic of incubation, which indicates that the effects of λ_3 will cancel out the overestimation part of λ_1. In this estimation, the interpretation is that the mandatory policy limits and decreased about 348 daily new cases (Robust SE = 15.46), 27720 cumulative cases (Robust SE = 1310.81), 16 daily deaths (Robust SE = 0.948), and 1867 cumulative deaths(Robust SE = 74.925) (all in per million people) compare with two other policies.

Table 2 indicates the DID estimation for group 2, without the data of no-control country. For group 2 estimation, the coefficients of $after_i$ and $mandatory : after$ become smaller, which verified that the optional policies are more effective than no-control policies. An interpretation of the estimation is that the country that implements mandatory policy will limit and decrease about 163 daily new cases (Robust SE = 8.49); 18008 cumulative cases (Robust SE = 770.89); 16 daily deaths (Robust SE = 1.35); 2026 cumulative cases (Robust SE = 103) (per million people) compare with the country that implements only optional policies after it was posted.

Table 3 indicates the DID estimation for group 3. Our interest of estimation is *optional* :

Table 2: Normal Difference in Difference model for second group

	Dependent variable:			
	‘New_cases/100‘	‘Cumulative_cases/100‘	‘New_deaths/100‘	‘Cumulative_deaths/100‘
	(1)	(2)	(3)	(4)
after	165.315*** (11.274)	18,573.560*** (1,009.235)	15.955*** (1.859)	2,053.680*** (136.971)
mandatory	−5.902 (31.339)	−294.082 (2,805.470)	−0.109 (5.168)	−29.531 (380.753)
mandatory:after	−162.966*** (33.222)	−18,008.530*** (2,974.043)	−15.823*** (5.478)	−2,026.181*** (403.631)
Constant	6.324 (9.528)	295.441 (852.960)	0.123 (1.571)	29.569 (115.762)
Observations	1,008	1,008	1,008	1,008
R^2	0.263	0.351	0.106	0.267
Adjusted R^2	0.261	0.349	0.104	0.265
Residual Std. Error (df = 1004)	140.033	12,535.890	23.092	1,701.347
F Statistic (df = 3; 1004)	119.370***	181.134***	39.865***	122.054***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Normal Difference in Difference for third group

	Dependent variable:			
	‘New_cases/100‘	‘Cumulative_cases/100‘	‘New_deaths/100‘	‘Cumulative_deaths/100‘
	(1)	(2)	(3)	(4)
after	633.598*** (32.902)	43,130.440*** (3,048.001)	17.847*** (2.323)	1,651.683*** (179.078)
optional	4.880 (35.879)	290.920 (3,323.779)	0.102 (2.533)	29.474 (195.280)
optional:after	−468.284*** (42.675)	−24,556.880*** (3,953.358)	−1.892 (3.013)	401.997* (232.270)
Constant	1.444 (27.563)	4.522 (2,553.398)	0.020 (1.946)	0.095 (150.018)
Observations	1,259	1,259	1,259	1,259
R^2	0.352	0.214	0.093	0.187
Adjusted R^2	0.351	0.212	0.091	0.185
Residual Std. Error (df = 1255)	337.577	31,272.610	23.836	1,837.344
F Statistic (df = 3; 1255)	227.372***	113.634***	42.901***	96.001***
Note:				*p<0.1; **p<0.05; ***p<0.01

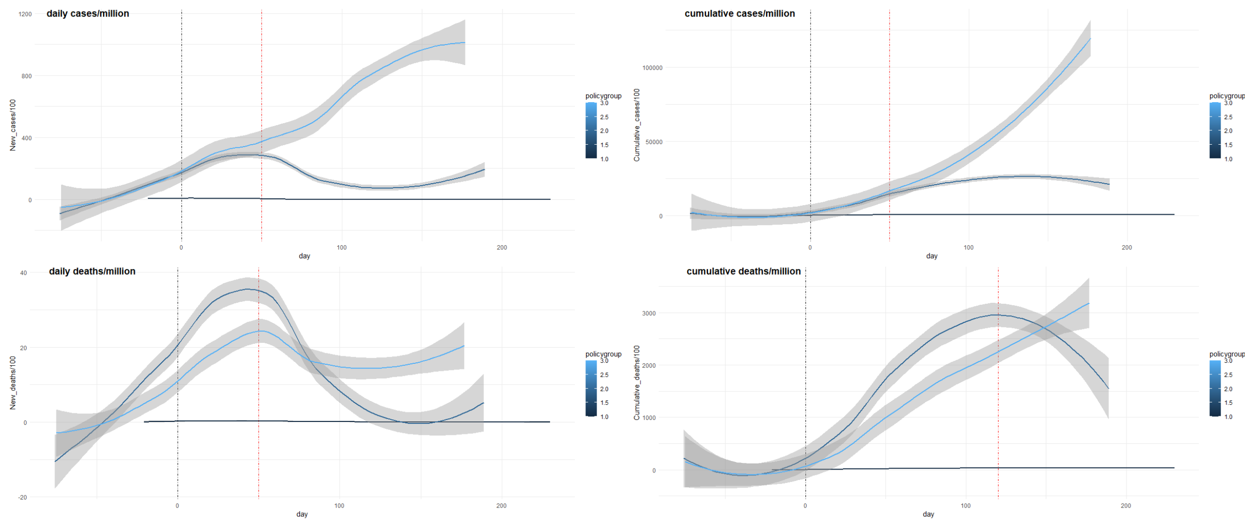

Figure 3: DID Line trends

$after$, which indicates that the country that implements optional policies will decrease and limit approximately 468 daily new cases (Robust SE = 32.67); 24557 cumulative cases (Robust SE = 3037.5); 2 daily deaths (Robust SE = 1.81) (All in per million people). However, it will increase about 402 cumulative deaths (Robust SE = 2.74) per million people compared with no-control policies. A possible explanation is that there are some factors that did not captured by the estimations, such as medical resources or universal health care. Note that the ψ_3 is statistically insignificant for daily deaths per million people with 70% confidence and only 90% confidence for cumulative deaths per million people.

Figure 3 indicates the line trends graph of the comparison between different policy groups for four dependent variables. The grey area represents the fluctuation of the data. Day 0 represents the last day of the pre-treatment and represents the boundaries of $after$. However, since the COVID-19 has an incubation period, I manually adjusted the vertical line to the day it took effect. For daily cases, daily deaths, and cumulative cases, I adjusted it to day 50 by considering the incubation period and lingering effects. For cumulative deaths, I adjusted it to day 120 since the medical help is able to extend the death period. Also, note that since its dependent variables are too small and stable for the mandatory-policy country, they were displayed on the area near X-axis, although it is hard to capture. As we can see in the figure 3 that mandatory-policy country's trends are quite stable, which they nearly all lies on the x-axis with little fluctuations. We see that they had little differences with no-control countries for optional-policy countries before treatments were posted. However, there are a large number of decreases in four dependent variables for optional-policy countries after their policies took effects. This indicates that compared with no-control policy, the optional policy can indeed limit the transmission and deaths within the countries. In contrast, in the figure 3, the dependent variables for no-control countries rise along with the time, indicating that the COVID-19 transmission and damages keep increasing through time. From the above interpretation, it is reasonable to conclude that under the COVID-19 situation, mandatory policies are highly effective in controlling and limiting the COVID-19 transmission and damages. Then, followed by optional policies, the no-control policies work least effectively.

5.2 Modified DID approach

Table 4 indicates the estimation results for the modified DID model. The model utilizes the *Country* factors to replace the treatment factor, which allows me to control the time in-variant country-specific factor so that the result will be less biased than the standard DID model. As we can see in 4, one of the estimation interests, $mandatory * after$, has more significant impacts on new cases, new deaths, and cumulative cases than the normal

Table 4: Modified Difference in Difference

	Dependent variable:			
	‘New_cases/100‘	‘Cumulative_cases/100‘	‘New_deaths/100‘	‘Cumulative_deaths/100‘
	(1)	(2)	(3)	(4)
China	107.397* (55.991)	13,277.330** (5,432.582)	12.326*** (4.522)	1,603.966*** (277.581)
India	−202.285*** (30.919)	−11,744.130*** (2,999.966)	2.109 (2.497)	559.312*** (153.285)
Canada	123.129*** (21.668)	14,607.960*** (2,102.315)	9.723*** (1.750)	1,175.959*** (107.419)
UK	197.143*** (21.706)	23,740.620*** (2,106.086)	25.178*** (1.753)	3,402.390*** (107.612)
US	423.247*** (30.969)	38,638.720*** (3,004.835)	22.693*** (2.501)	2,662.663*** (153.534)
Japan				
mandatory:after	−628.302*** (59.175)	−42,327.240*** (5,741.547)	−17.617*** (4.779)	−1,614.241*** (293.368)
optional:after	−456.011*** (30.726)	−23,195.360*** (2,981.205)	−0.649 (2.481)	565.603*** (152.327)
after	630.651*** (23.638)	42,892.270*** (2,293.502)	17.750*** (1.909)	1,641.740*** (117.188)
Constant	−106.975*** (21.188)	−13,275.970*** (2,055.832)	−12.312*** (1.711)	−1,603.928*** (105.044)
Observations	1,511	1,511	1,511	1,511
R^2	0.621	0.491	0.290	0.609
Adjusted R^2	0.619	0.489	0.286	0.607
Residual Std. Error (df = 1502)	243.090	23,586.080	19.632	1,205.146
F Statistic (df = 8; 1502)	307.087***	181.300***	76.735***	292.847***
Note:				*p<0.1; **p<0.05; ***p<0.01

DID model (group 1). However, the cumulative deaths coefficients became higher than the standard DID model. A possible interpretation of such results is that the $Country$ variables contain some time in-variant factors such as medical resources, universal health care, or infrastructures capabilities, the standard DID model did not capture the time in-variant factors. Same with the $optional * after$ coefficients, the differences are quite small, but $Country$ variables contain factors that the normal DID model didn't capture. Note that although Japan's estimation is fully blanked, the *constant* coefficients contain the estimations of Japan's data. Nearly all the estimations are statistically significant, but the *newdeaths* estimations are statistically insignificant for variable $India$ and $optional * after$, this may be due to the issues of data itself.

6 Conclusion

Early policies should intend to decrease and limit the transmission of COVID-19. By observing and analyzing the data, it is reasonable to conclude that mandatory policy is highly correlated with improving the COVID-19 situation. Optional policies, compared with no-control policies, somehow improved the COVID-19 situation. Despite the estimation that it decreases 468 daily new cases (with Robust SE = 32.67) and 24557 cumulative cases (with Robust SE = 3037.5), its effects on daily deaths is statistically insignificant as well as it will also increase the 402 cumulative deaths (Robust SE = 2.74), this may due to the omitted variables such as medical resources or universal health cares. Figure 3 also proved the above estimations. In conclusion, Mandatory policies are the best way to stop and limit the transmission and damages of COVID-19, followed by optional policies. Lastly, no-control policies do nothing to the COVID-19. However, due to the different aspects and points of view, some people may argue from the economic aspect or human rights aspect that mandatory policy, or even optional policies, are highly impossible to achieve. The truth is that COVID-19 damages are larger with less control, I found in my past analysis that the economic issues raised by COVID-19 are due to the ripple effects that led by COVID-19 transmission to a large extent, such as liquidity trap and bankruptcy of small and medium-sized enterprises, what I want to say is that the faster the control, the less the damages. Back to the question, should government officials take responsibility for those infected or killed by COVID-19? Yes, and this is only my answer.

References

[CDC22] United States of America CDC. Excess deaths associated with covid-19, Mar 2022.

[HAG+21] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 5(4):529–538, Apr 2021.

[LW20] Wei Lyu and George L. Wehby. Community use of face masks and covid-19: Evidence from a natural experiment of state mandates in the us. *Health Affairs*, 39(8):1419–1425, 2020. PMID: 32543923.

[Moo20] Imad A. Moosa. The effectiveness of social distancing in containing covid-19. *Applied Economics*, 52(58):6292–6305, 2020.

[RMRG+20] Hannah Ritchie, Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian, and Max Roser. Coronavirus pandemic (covid-19). *Our World in Data*, 2020. https://ourworldindata.org/coronavirus.

[TLL+20] Huaiyu Tian, Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz U.G. Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, Ben Wang, Peng Yang, Yujun Cui, Yimeng Song, Pai Zheng, Quanyi Wang, Ottar N. Bjornstad, Ruifu Yang, Bryan T. Grenfell, Oliver G. Pybus, and Christopher Dye. The impact of transmission control measures during the first 50 days of the covid-19 epidemic in china. *medRxiv*, 2020.