The Impact of Rapid State Policy Response on Cumulative Deaths Caused by COVID-19

Takuto Nakamura
Cambridge School of Weston; Massachusetts, USA
Email: tnakamura2021@csw.org

Abstract – Amidst the COVID-19 pandemic the majority of US states enacted a statewide stay-at-home order. This paper makes use of state-level heterogeneity in policy responses to conduct a cross-sectional regression analysis of lock-down response on Covid-19 deaths. The results of this paper suggest that states which imposed stay-at-home orders more quickly after reaching a threshold proportion of deaths had significantly fewer deaths 3 weeks after lock-down. We use of a robust set of control variables in order to minimise bias arising from confounding factors. Results show that the predicted increase in the cumulative deaths after 21 days of lockdown for the average state associated with delaying implementation of policy by 1 week is approximately 46%. The result shows that speed is a crucial factor that state governors should consider when enacting a stay-at-home policy.

Key Words – COVID-19, Lockdown speed, Stay-at-home, Ordinary least squares, Cumulative deaths

INTRODUCTION

I. *Background*

COVID-19 is an infectious disease caused by the SARS-CoV-2 virus that commonly affects the respiratory system and has been classified as a pandemic by the World Health Organisation [1]. In the United States the federal political system grants each state the authority to enact policies at their discretion. This has caused statewide responses to contain the virus to be enacted at different times; policies include stay-at-home orders, face-covering requirements, large gathering bans, and travel restrictions. Measuring the effectiveness of policies is commonly done by tracking cases and deaths. In this paper, we chose to look solely at deaths. Stay-at-home orders have been a popular measure for states to enact. Stay-at-home orders mandate citizens to restrict their movement outside of their residence. Given the prominence of this policy, a number of papers have already analysed 2020 stay-at-home orders.

II. *Literature Review*

Tian et al. [2] looked at the effect of travel restrictions and other control measures on the prevention of the spread of SARS-CoV-2 in Wuhan, China in the first 50 days of the pandemic. The team used multiple methods of analysis including ordinary least squares (OLS) estimation. Results show that a faster response in intracity transports suspension, entertainment venue closures, and public gathering bans lower the number of COVID-19 cases in the first week. Their model suggested that the national emergency response and travel bans helped limit the confirmed COVID-19 cases by 96% compared to a model where none of these measures were taken into effect.

Fowler et al. [3] researched the effectiveness of stay-at-home orders at a county level in the US. Using data from the COVID Tracking Project, the team graphed the log of confirmed cases of multiple counties and analysed the correlation between the counties that implemented stay-at-home orders and ones that didn’t. The team found that, overall, counties that implemented stay-at-home orders saw a reduction in COVID-19 cases compared to counties that did not implement any policy. I used this paper as a basis for our model and analysis.

With the continued lack of containment of COVID-19 cases, we still do not have a clear understanding of measures that are effective in containing COVID-19. By analysing the effect of policies, we can prepare the most effective way of dealing with continuous waves of this current pandemic or a new pandemic in the future. Here, we aim to analyse the correlation between the speed of lockdown and deaths caused by COVID-19.

DATA

Key data used in the analysis comes from the Kaiser Family Foundation[4], a non-profit organisation mainly focusing on medicare and health policies. Population data is from the 2010 and 2019 census [5], income is from the 2016 American community survey [6], and political affiliation is from the 2016 presidential elections [7].

Starting this project our main goal was to look at the effects of policies on the number of deaths that occurred after the policies are enacted. In doing this we first looked at the different policies states implemented during this pandemic. A few common policies were stay-at-home orders, face-covering requirements, gathering bans, and travel restrictions. We chose to look at the stay-at-home policy for our project because this was a policy most states put out and in our hypothesis we thought had the most effect out of the other common policies. From the paper mentioned above, we

TABLE 1: Descriptive Statistics of Key Variables

VARIABLES	Mean	Median	p25	p75	S.D.	#Obs	min	max
death21days	406.0	183	87	537	575.5	41	7	2,854
lockdownspeed	2.341	2	-2	5	6.147	41	-10	21
income	61,141	58,756	54,478	70,315	10,337	41	44,097	83,242
politics	0.561	1	0	1	0.502	41	0	1
population	7.210e+06	5.640e+06	2.976e+06	8.536e+06	7.557e+06	41	623,989	3.951e+07
travelquarantine	0.512	1	0	1	0.506	41	0	1
facecovering	2.439	3	2	3	0.976	41	0	3
urbanpopulation	74.60	75.10	66.30	87.20	14.84	41	38.70	95

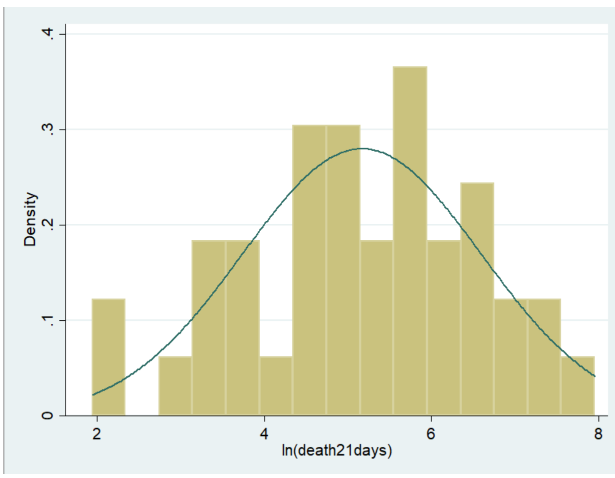

FIGURE 1B: Histogram of logarithm of death21day excluding New York

know that there was a significant effect of having a lockdown at the county level. Our goal is to see if the speed at which this policy is implemented has an effect on the outcome. To measure the speed of lockdown we first thought of choosing a single date as a benchmark to see the difference in timing but this would cause bias in states with earlier COVID-19 cases. Based on existing analyses [8] we came up with taking the number of days between the day at which 1 in 1 million deaths was reported in the state and the day the stay-at-home policy was implemented. This allowed us to get a more reliable benchmark for the speed of lock down because this would adjust for the timing at which each state was introduced with COVID-19 and the differences in population for each state.

In order to evaluate the effect of the stay-at-home order, we looked at the number of deaths that happened after the policy. We chose to solely look at the deaths because of differences in the number of tests in different states, do not allow for a like for like comparison of cases across the states. If early policies were more effective we should see a lower number of deaths in such states.

In order to have a benchmark for the deaths, we follow Fowler et al. [3] in looking at the day in which the lockdown was initiated and taking the number of cumulative deaths 3 weeks (21day) after the enactment of the policy. We believe 21 days is a sufficient period of time to capture most of the deaths prevented due to lockdown based on calculations on the average incubation period and median lag time for fatality.[1] This allows like for like comparison for each state.

In our final data set, we only looked at the mainland state which had a lockdown policy implemented at some point. This was to get a more similar comparison between states. We also excluded New York from our data set because of its anomalous nature. (Figure 1A) This then left us with 41 states in our data set.

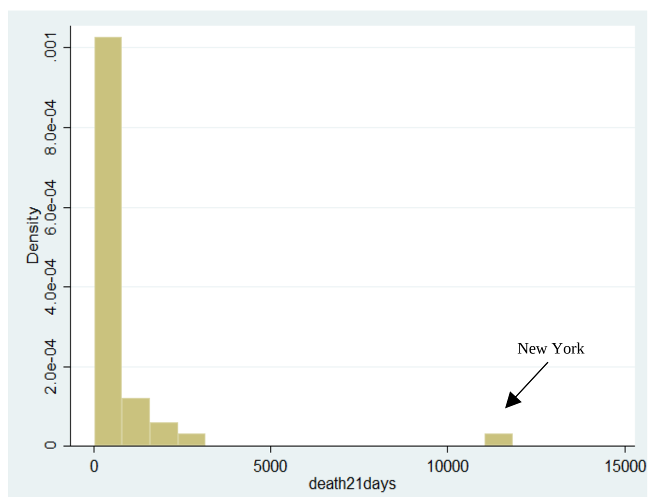

FIGURE 1A: Histogram of death21days

[1]Based on the data from the WHO and the CDC the average incubation period is 5-6 days and the median lag time for fatality is 13 days which, assuming the lockdown has an immediate effect on transmission implies that most of the prevented deaths would have occurred 18-19 days after lockdown. [9][10]

TABLE 2: Variable Descriptions

Variable	Definition	Purpose
lockdownspeed	The number of days between 1 in 1 million COVID-19 deaths and implementation of a lockdown policy. (Higher values mean slower response time)	Independent Variable
death21days	The number of cumulative COVID-19 deaths 21 days after the implementation of a lockdown.	Dependent Variable
Politics	Republican or democratic state taken from the 2016 presidential election.	Control Variable
Urbanpopulation	% of population classified as urban taken from the 2010 census.	Control Variable
Population	State populations were taken from the 2019 census.	Control Variable
Income	Average state median household income was taken from the 2018 American community survey.	Control Variable
Travel Quarantine	Dummy variable describing the occurrence of a travel quarantine.	Control Variable
Face Covering	Categorical variable describing the different strengths of a face covering policy. Required for General Public (3) Required for Certain Employees (2) Allows Local Officials to Require for General Public (1)	Control Variable

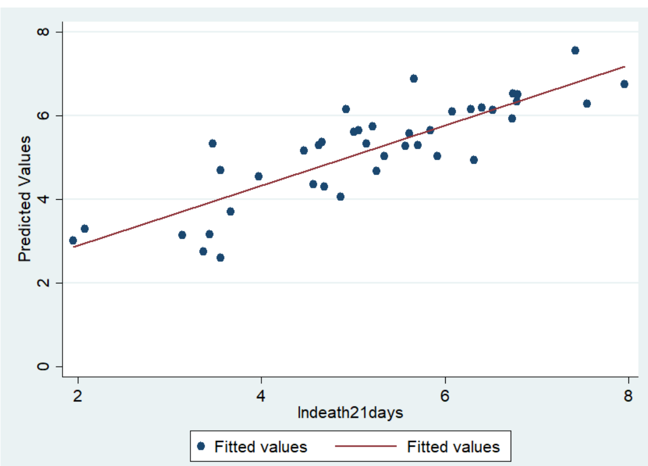

FIGURE 2: Scatter plot of predicted values on the log of number of deaths 21 days after lockdown.

METHODOLOGY

$$\text{lndeath21days}_i = \beta_0 + \beta_1 lockdownspeed_i + \boldsymbol{X_i}\,\boldsymbol{\gamma} + \varepsilon_i$$

The above model is what we sought to estimate with the coefficient of interest being beta 1. The beta 1 coefficient multiplied by 100 is interpreted as the average predicted percentage point difference in the number of cumulative deaths after 21 days in a given state, after controlling for state differences, if the lockdown policy was delayed by 1 day. We considered a quadratic term on our independent variable because of the exponential nature of the spread of viruses. However, in this more flexible specification, we found the coefficient on the square term to be negative and the p-value to be insignificant indicating the linear model to be more appropriate. This is most likely because the early stages of a pandemic are best described in a linear fashion [11].

We estimate the above model using OLS regression. OLS takes the squared difference in the dependent variable and the predicted linear model and minimises the sum of the differences in order to estimate the relationship of the dependent variable on the independent variable. To obtain unbiased and efficient estimates, least squares has a number of assumptions that need to be met. These are discussed below.

In order to demonstrate the normality of residuals, we applied the Shapiro-Wilk test [12] to our residuals. We realised that the dependent variable is log normally distributed through graphical analysis. (Figure 1B)

Hence we used the natural log of the dependent variable in our model. Furthermore, a Shapiro-Wilk test performed on the residuals gives us a p-value of 0.128 which is greater than 0.05, meaning that we fail to reject the null hypothesis of the residuals being normally distributed.

To deal with heteroscedastic terms we used robust standard errors.

To ensure that multicollinearity does not have a significant effect on the standard errors of the estimated coefficients we performed a VIF test, the result of which was

TABLE 3: OLS Results				
VARIABLES	(1)	(2)	(3)	(4)
lockdownspeed	0.0519*	0.0730***	0.0807***	0.0660***
	(0.0291)	(0.0238)	(0.0203)	(0.0224)
population		1.01e-07***	3.94e-08	2.60e-08
		(3.11e-08)	(2.51e-08)	(2.59e-08)
politics			0.372	0.620
			(0.389)	(0.410)
income			-3.98e-06	-6.31e-06
			(2.39e-05)	(2.74e-05)
urbanpopulation			0.0613***	0.0647***
			(0.0142)	(0.0151)
travelquarantine				-0.799**
				(0.303)
1.facecovering				-0.251
				(0.418)
2.facecovering				-0.202
				(0.502)
3.facecovering				0.118
				(0.372)
Region Controls?	N	Y	Y	Y
Observations	41	41	41	41
R-squared	0.050	0.435	0.673	0.743

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

1.92; this low value [13] provides sufficient evidence that we do not have significant multicollinearity in our model.

Strict exogeneity is required in order for the least squares estimator to be unbiased. This can be caused by multiple different factors such as measurement error in the independent variable, simultaneity, and omitted variables. It is reasonable to assume that there is almost no measurement error in the independent variable because we know the exact dates in which the policies were taken into effect. Simultaneity can also be ruled out because our values of the dependent variable were realised 21 days after our independent variable. The most likely source of bias is from omitted variable bias. In our model, we included a range of control variables to minimise omitted variable bias.

In the table above we describe each variable and the purpose it has in our model. (Table 2)

RESULTS

As expected the simple regression model in column 1 (Table 3) shows a positive correlation between speed of lockdown and the cumulative deaths after 21 days but is not statistically significant. However, this result fails to take into account potentially confounding variables. The 2nd column includes population size and region controls which increases the average predicted impact of delayed policy implementation on the cumulative deaths after 21 days of lockdown and the result is now statistically significant at the 1% level. The 3rd column includes additional controls such as urban population, income, and politics demonstrating the robustness of the result: the coefficient on lockdownspeed remains positive and statistically significant. Finally in the 4th column we take into account some of the other policies that were implemented in a similar timeframe to ensure we are not capturing the effect of policies other than lockdown. The magnitude of the coefficient decreased slightly but is still significant at the 1% level.

From our results, we can see that the coefficient does not change greatly in the different variations of the model ranging from 5-8 percentage points and that the p-value of the coefficient implies that there is a statistically significant relationship.

The result from our preferred specification (column 4) suggests that the expected increase in the cumulative deaths after 21 days of lockdown for the average state associated with delaying implementation of policy by 1 week is 46.2%. With the average number of COVID-19 deaths after 21 days of lockdown being 408 deaths, if we imagine a scenario in which the “average” state implements a stay-at-home policy a week later, we would expect to see an increase in deaths from 408 to 596.

DISCUSSION & LIMITATIONS

These results strongly support rapid action from state governments in response to the spread of COVID-19. Additionally as seen in the table urban population and travel quarantine seem to also have a significant effect on the deaths. As expected the urban population has a positive coefficient and travel quarantine has a negative coefficient. This shows that higher urban population % is correlated with an increase in cumulative deaths after 21 days of lockdown and that a travel quarantine is associated with lower deaths.

The R-squared value shows that our preferred specification has high predictive power. Furthermore, this is represented in figure 2 with our predicted deaths strongly correlating with the actual number of deaths.

While we were able to do an analysis on a data set with 41 observations this is rather small from a statistical standpoint and consequently, the power of the statistical tests conducted is low. In future improvements to this analysis, it would be better if we could gather more data. For example, we can work on a county scale, not state-level as done by Fowler et al. [3].

Although we attempted to account for confounding factors, we cannot rule out the possibility of endogeneity and therefore bias as we don't account for every variation between states, such as the age distribution of the state. As such we cannot say with certainty that our results are causal. One possible variable that may affect our data is the availability of beds for COVID-19 patients in each state. This theoretically could have caused a faster response to COVID-19 as the state has a smaller capacity to take in patients and will cause greater deaths after 21days because of the lack of medical preparation the state had. Assuming this is plausible our failure to account for this in our analysis implies our estimate of the impact of the speed of lockdown is an underestimate. Further analysis needs to be conducted.

In our paper, we looked at the stay-at-home policy to be the most effective policy when it came to COVID-19 response. But it is true that there were several other policies that were enacted within days of each other. Although we account for some of these, we cannot say that our findings capture solely the effect of the speed of lockdown but instead captures the effect of lockdown and other bundled policies enacted in the same timeframe.

CONCLUSION

We analysed the effect of the speed of lockdown on the number of deaths after 21 days of lockdown. This shows that the speed of lockdown is a crucial factor that state governors should consider. However, we recommend policymakers take in to account other research of other policies because we do not have a measure to compare different aspects of multiple policies; there is a possibility that there are other factors more impactful than the speed in which the lockdown is enacted. Furthermore, we are not comparing the magnitude of the effect of the speed of lockdown to other significant policy measures like the strength of lockdown, the duration of lockdown, and other policies. We leave this to further research.

AKNOWLEDGEMENTS

The author wishes to thank Matin Khadem for his help and guidance in conducting research and proofreading this paper.

REFERENCES

[1] World Health Organization: WHO. "Coronavirus." World Health Organization, 10 Jan. 2020, www.who.int/health-topics/coronavirus#tab=tab_1.

[2] Tian, Huaiyu, et al. "An Investigation of Transmission Control Measures during the First 50 Days of the COVID-19 Epidemic in China." Science, vol. 368, no. 6491, 2020, pp. 638–42. Crossref, doi:10.1126/science.abb6105.

[3] Fowler, James H., et al. "The Effect of Stay-at-Home Orders on COVID-19 Infections in the United States." SSRN Electronic Journal, 2020. Crossref, doi:10.2139/ssrn.3576826.

[4] "State Data and Policy Actions to Address Coronavirus." KFF, 25 Aug. 2020, www.kff.org/coronavirus-covid-19/issue-brief/state-data-and-policy-actions-to-address-coronavirus.

[5] US Census Bureau. "State Population Totals: 2010-2019." The United States Census Bureau, 30 Dec. 2019, www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html.

[6] ---. "Data Profiles." US Census Bureau, www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2016. Accessed 26 Aug. 2020.

[7] "Associated Press Interactive." Associated Press, interactives.ap.org/2016/general-election/?SITE=APQA. Accessed 26 Aug. 2020.

[8] "Coronavirus: The Lost Six Weeks When US Failed to Contain Outbreak." BBC News, 12 May 2020, www.bbc.com/news/av/world-us-canada-52622037.

[9] "Coronavirus disease 2019 (COVID-19) Situation Report – 73" World Health Organization ,2 April 2020, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf

[10] "Case-Fatality Risk Estimates for COVID-19 Calculated by Using a Lag Time for Fatality - Volume 26, Number 6-June 2020 - Emerging Infectious Diseases Journal - CDC." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, wwwnc.cdc.gov/eid/article/26/6/20-0320_article.

[11] Viboud, Cécile, et al. "A Generalized-Growth Model to Characterize the Early Ascending Phase of Infectious Disease Outbreaks." Epidemics, vol. 15, 2016, pp. 27–37. Crossref, doi:10.1016/j.epidem.2016.01.002.

[12] Shapiro, S. S., and M. B. Wilk. "An Analysis of Variance Test for Normality (Complete Samples)." Biometrika, vol. 52, no. 3/4, 1965, p. 591. Crossref, doi:10.2307/2333709.

[13] Daoud, Jamal I. "Multicollinearity and Regression Analysis." Journal of Physics: Conference Series, vol. 949, 2017, p. 012009. Crossref, doi:10.1088/1742-6596/949/1/012009.