

Predicting COVID-19 Unreported Case From Space

Yayan Satyakti^{1*}

¹ Department of Economics, Faculty of Economics and Business, University of Padjadjaran, Bandung

* yayan@unpad.ac.id

DOI
10.30536/jkkpa.v2n1.5

Abstract

This paper investigates how social behavior motives such as people's mobility responding to the news, containment policy, and spreading positive case of COVID-19. I assumed that the government conduct news, and containment policy to reduce human mobility as well as spreading new cases. In addition, I assumed that household behavior from medium or higher income group that located in an urban area has a significant effect on increasing mobility and spreading of new cases. I proxied, night lights data that indicates the geographic economic performance of a particular income group is the source of mobility behavior for commuting and traveling. The mobility data retrieved from Waze Car Congestion Report according to an hourly basis that was located with spatial GPS. The confirmed case of COVID-19 was retrieved at the provincial level. The prediction of the confirmed case model was estimated with the econometric epidemiologist model. The results indicate that without policy the spreading case of COVID-19 will increase about 8 million cases, whereas with policy the unreported COVID-19 confirmed cases about 1.9 – 2 times of government publicly report. We conclude that improving literacy and awareness of COVID-19 is necessary to be important agenda rather than an economy. The results indicate that Indonesian prefer to engaging with economic rather than health policy.

Keywords: COVID-19, econometric method, human mobility

1. Introduction

Since early January 2020, around the world has devastated how COVID-19 impacts the global economy. Nevertheless, human mobility is a prominent factor that affects on spreading of confirmed cases. In the context of this study, several studies such as Hsiang et al. (2020), Gatto et al. (2020), or Hanming et al. (2020) have been studied how human mobility and containment policy affect on spreading case of COVID-19. In an earlier previous study such as Charu et al. (2017) have studied how the human mobility effect of influence pandemic in the USA. In the Indonesia context such as Satyakti (2020) evaluate with regression discontinuity approach how human mobility effect on spreading confirmed case of COVID-19. The studies stated that restriction of human mobility effects of reducing the spread of COVID-19 confirmed the case.

Indonesia at this moment has become a crucial COVID-19 hotspot confirmed case in the South East Asian countries. Foreign press such as The Sydney Morning Herald stated Indonesia will become the third country after China, India as a peak hotspot of COVID-19. According to Satyakti (2020), one of the crucial factors that affect human mobility is containment policy, law enforcement, human literacy, and social behavior of COVID-19 awareness. Lack of law enforcement about Health Protocol and Human Mobility is a crucial issue that affects on spreading of COVID-19 confirmed case. While the previous study such as Fogli and Veldkamp (2020)

strongly stated that social behavior effect on increasing spreading of pandemics. When the policymaker unable to enforce the social behavior to reduce human mobility. If this policy unable to achieve the confirmed case still spreading around society.

If we looking at the previous studies, the model usually evaluates human mobility in aggregate ways without concerning social behavior in terms of spatial economic performance that affects human mobility. In the previous study such as Charu et al. (2017), Lemey et al. (2014) in epidemiology stated that human mobility especially in the urban area strongly interacts and improving significantly pandemics. Unfortunately in a pandemic model, there is no determining variable of how human mobility affects by economic behavior. In fact, in Gatto et al., (2020) model. The model is not included how economic behavior affects human mobility. On the other hand, Glaeser et al. (1992) stated that human mobility affects by economic behavior and agglomeration. In a recent paper as Çolak et al., (2016) stated that human mobility and economic behavior are strongly connected and increase economic performance to some extent. In terms of these issues, we need further to investigates more by providing new variables to assess this issue.

Generally, measuring economic performance with conventional data is challenging and requires a huge budget and time lag. In several years, many economists have proposed various methods to tackle these issues. The promising method is light night data that usually capture the economic performance spatially. Henderson et al. (2012) have proposed this method how to estimate economic performance from satellite data such as light nights. The advantage of this data, we can predict in more detail how economic performance is assessed with satellite data to perform spatial economic performance into more detailed regions that usually conventional data unable to capture these data unavailability. From this standpoint, we can state that evaluating social behavior according to economic performance with satellite data is doable to perform.

This paper aims to investigate the effect of social behavior through spatial economic performance on predicting unconfirmed cases. I choose this topic because there several caveats of public data due to lack of testing about the confirmed case. The previous study such as Hsiang et al., (2020) for the global case, and Hanming et al., (2020) reported how public data in China is uncertain to measure the confirmed case. My paper has endeavored to contribute to this issue. Hence, we can figure out how spatial performance affects human mobility and spreading the COVID-19 confirmed case.

2. Data

I combined three sources of data that is car traffic congestion data, map data, and COVID-19 daily case data. Car congestion traffic was obtained from Waze Dashboard Data. This data is not publicly available, we should become a partner to retrieve the data. The Waze provides detailed information such as hours report (i.e. especially for jam report that consists standstill traffic, heavy traffic, and moderate traffic). In hours report Waze also reports car flow of origin and destination to a specific destination. While map data obtained from the Indonesian Geospatial Agency with 1.000.000 scale with Earth Mercator projection.

I employed confirmed case data according to province level. The reason why I used the data because this data provides consistent data availability in terms of daily frequency at the province level across Indonesia. This is the only data published by the government, officially.

To support spatial mobility across space and time. I mapped human mobility into detail location car congestion mobility within a sub-district level. I should reduce mobility bias that some of the regions such as West Java Province have more mobility than other regions because it is crossed by highways. I should calculate, which car

on highways will go out, stop and stay in particular regions outside highways mobility. Otherwise, if we estimate the spatial mobility aggregated with larger boundaries such as district, it will increase bias estimation between car congestion and spreading of COVID-19 case. Therefore, it is necessary when the unit analysis has a deeper level such as sub-level district level. We can map human mobility across spatial.

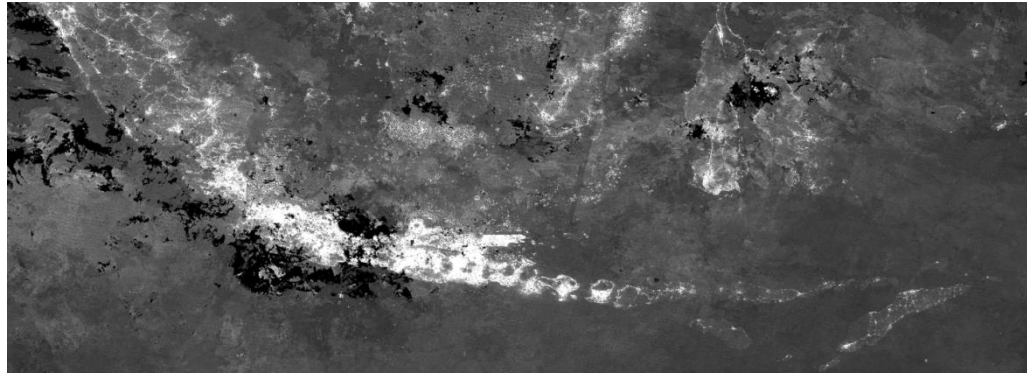


Figure 1. Night Lights Data (in March 2020)

Source: https://eogdata.mines.edu/download_dnb_composites.html

To get mobility flow between region, for each observation, I mapped the car GPS location that provides by Waze. So, I can impose on the map for an hourly report with the GIS software. On other hand, the car located will produce bias interpretation whether they in taxing or stay at the end of the journey. To reduce this bias, the data was collected with aggregated daily data to ensure that one region has a more consistent flow of car congestion data rather than hourly data. This method has produced consistent panel data set in time with daily frequency as well as spatial boundaries for each sub-level district (called *desa*) across Indonesia. The night light dataset was retrieved from the Earth Observation Group (EOG). This site has been provided consistently in terms of projection and satellite datasets across the world. I focus on the Java area due to the highest COVID-19 confirmed case across Indonesia.

3. Methods

I adopted an epidemic model according to Charu et al. (2017) that increasing incidents of COVID-19 in one location determined by works commutes and other transportation means that connect social interaction. The equation consist of :

$$\ln \ln C_{rt} = \alpha_0 + \alpha_1 \ln \ln V_{rt} + \alpha_2 DSRV_{rst} + \alpha_3 DT_{irt} + \eta_{rt} + \epsilon_{rt} \dots\dots\dots(1)$$

Where α_0 s constant in the epidemic model, r is a sub-district at the lowest level of administrative boundaries, is sub-district within the center of the infected area. C is an additional infected case in person at the province level, and V is human mobility that proxied by work commutes that indicates traffic report congestion with the number of car reports in each hour. DT is a news policy that affects human mobility.

While the data were only available for March 2020. We should portray how this social behavior data demonstrate human mobility activity. To deal with this issue, because the income group affected in a longer period rather than a short period. The dynamic effect will be less effective, they were unable to sell their assets so the location such as a house or other fixed effects will still similar to the previous period. Otherwise, to portrayed whether the level of income group either stay or

mobile and follow the human restriction policy. We interact the mobility data with pixel-level of nightlight data through the interaction of $DSRV_{rst}$. We can predict this interaction with parameter α_2 . DSRV is variable interaction between human mobility and spatial economic performance.

The ideal parameter for this variable should be negative. The mobility should be decreasing in higher income levels. I assumed that higher level along with higher education level (Fields, 1975). In emerging countries like Indonesia, I assumed higher income levels will have more educated, more awareness, and better health behavior. Otherwise, the results are the opposite, I can identify that better income level in this location unaware of the health policy of COVID-19.

The η_{rt} is sub-district time fixed effect and ϵ_{rt} is clustered by sub-district and time. This model is a general form of fixed effect econometric model that indicates how additional of COVID-19's case determined by mobility and treatment effect for both spatial and time in sub-district area. Since the sub-district treats as a fixed effect and some of the areas were empty. We estimate equation (1) with a fixed effect with a large dummy variable.

4. Results

We depicted human mobility that proxy with car congestion report during March 2020 until the end of May 2020 in Table 2 and Table 3. The results indicate that there is significant decreasing mobility during March 2020, April 2020, and May 2020. The illumination lights in May 2020 are lighter than in March 2020. The night lights were plotted with pixel polygon that represents the lowest spatial level that portrayed household income level. When the pixel is lighter than others, the pixel has more energy consumption, more electrical appliances, and more income, otherwise darker vice versa. As we noted earlier this data should be confirmed with mobility data that proxy by car congestion report. When some pixel is lighter and has more mobility which means that the region has a higher level of income and more mobility.

I can confirm there is a positive relationship between Human Mobility and Light Night Times across sub-district levels in Indonesia. The sub-district level which lighter or richer has more mobility rather than darker. There is a supportive behavior that improving income increase mobility behavior. Another evidence that I am going to describe is whether the policy's effect on reducing behavior and confirmed the case of COVID-19. In Table 1, I present an estimation of everyday cases and the important news policy announcement. The results indicate that a significant reduction has existed when the first case was announced. On the other hand, human mobility increases when the human restriction mobility policy announce such as on Working From Home (WFH) policy. The progress of human mobility is decreasing along with the confirmed case news and shock policy such as Regional Emergency Status that announced in specific provinces e.g. DKI Jakarta and West Java.

In particular of Human Mobility Parameter, there is a significant decrease for each news on reducing COVID-19 confirmed case. So the policy is effective to brought people awareness for COVID-19 spread. On the other hand, if we looking at the interaction variable between Night Light and Human Mobility Report. The parameter has the opposite sign. The people with higher income level regions and more mobility increase their mobility. The policy news was unable to reduce their mobility behavior.

Table 1: News and Mobility
 Dependent variable: growth rate of active cases

VARIABLES	(1) Coefficient	(2) Std Error	(3) P-value
Ln(Car Congestion Report)	-0.066***	0.011	0.000
Ln(Night LightsxHum. Mob. Report)	0.039***	0.004	0.000
2-Mar-20 – First Case	-8.330***	0.014	0.000
3-Mar-20 – Confirm. Case. Rep.	-8.325***	0.014	0.000
4-Mar-20 – Confirm. Case. Rep.	-8.331***	0.014	0.000
5-Mar-20 – Confirm. Case. Rep.	-8.335***	0.014	0.000
6-Mar-20 – Confirm. Case. Rep.	-7.641***	0.014	0.000
7-Mar-20 – Weekend	-7.616***	0.013	0.000
8-Mar-20 – Weekend	-7.190***	0.012	0.000
9-Mar-20 – Confirm. Case. Rep.	-6.069***	0.014	0.000
10-Mar-20 – Confirm. Case. Rep.	-5.721***	0.014	0.000
11-Mar-20 – Confirm. Case. Rep.	-5.489***	0.014	0.000
12-Mar-20 – Confirm. Case. Rep.	-5.496***	0.014	0.000
13-Mar-20 – Confirm. Case. Rep.	-4.786***	0.014	0.000
14-Mar-20 – Weekend	-4.418***	0.012	0.000
15-Mar-20 – WFH Announced	-5.277***	0.037	0.000
16-Mar-20 – Confirm. Case. Rep.	-4.840***	0.029	0.000
17-Mar-20 – Confirm. Case. Rep.	-4.498***	0.034	0.000
18-Mar-20 – Confirm. Case. Rep.	-4.125***	0.023	0.000
19-Mar-20 – Reg. Emergency Status	-3.861***	0.025	0.000
20-Mar-20 – Confirm. Case. Rep.	-3.630***	0.022	0.000
21-Mar-20 – Weekend	-3.387***	0.020	0.000
22-Mar-20 – Weekend	-3.200***	0.022	0.000
23-Mar-20 – Confirm. Case. Rep.	-3.115***	0.020	0.000
24-Mar-20 – Confirm. Case. Rep.	-3.005***	0.023	0.000
25-Mar-20 – First Rapid Test	-2.912***	0.026	0.000
26-Mar-20 – Confirm. Case. Rep.	-2.803***	0.020	0.000
27-Mar-20 – Confirm. Case. Rep.	-2.613***	0.019	0.000
28-Mar-20 – Weekend	-2.492***	0.021	0.000
29-Mar-20 – Weekend	-2.323***	0.020	0.000
30-Mar-20 – Confirm. Case. Rep.	-2.207***	0.018	0.000
31-Mar-20 – Confirm. Case. Rep.	-2.169***	0.018	0.000
1-Apr-20 – Confirm. Case. Rep.	-2.082***	0.018	0.000
2-Apr-20 – Confirm. Case. Rep.	-2.067***	0.020	0.000
3-Apr-20 – Confirm. Case. Rep.	-1.952***	0.016	0.000
4-Apr-20 – Weekend	-1.908***	0.017	0.000
5-Apr-20 – Weekend	-1.841***	0.017	0.000
6-Apr-20 – Confirm. Case. Rep.	-1.775***	0.014	0.000
7-Apr-20 – Confirm. Case. Rep.	-1.631***	0.014	0.000
8-Apr-20 – Confirm. Case. Rep.	-1.601***	0.015	0.000
9-Apr-20 – Confirm. Case. Rep.	-1.509***	0.014	0.000

VARIABLES	(1) Coefficient	(2) Std Error	(3) P-value
10-Apr-20 – Confirm. Case. Rep.	-1.480***	0.017	0.000
11-Apr-20 – Weekend	-1.446***	0.016	0.000
12-Apr-20 – Weekend	-1.241***	0.014	0.000
13-Apr-20 – Confirm. Case. Rep.	-1.152***	0.012	0.000
14-Apr-20 – Confirm. Case. Rep.	-1.106***	0.012	0.000
15-Apr-20 – Jakarta PSBB Ann.	-1.064***	0.012	0.000
16-Apr-20 – Confirm. Case. Rep.	-1.015***	0.012	0.000
17-Apr-20 – Confirm. Case. Rep.	-0.966***	0.012	0.000
18-Apr-20 – Weekend	-0.917***	0.013	0.000
19-Apr-20 – Weekend	-0.873***	0.013	0.000
20-Apr-20 – Confirm. Case. Rep.	-0.838***	0.012	0.000
21-Apr-20 – Confirm. Case. Rep.	-0.797***	0.011	0.000
22-Apr-20 – PSBB Regions	-0.804***	0.012	0.000
23-Apr-20 – Confirm. Case. Rep.	-0.757***	0.013	0.000
24-Apr-20 – First Ramadhan	-0.727***	0.013	0.000
25-Apr-20 – Weekend	-0.684***	0.014	0.000
26-Apr-20 – Weekend	-0.671***	0.015	0.000
27-Apr-20 – Confirm. Case. Rep.	-0.648***	0.011	0.000
28-Apr-20 – Confirm. Case. Rep.	-0.580***	0.012	0.000
29-Apr-20 – Cont'd PSBB	-0.592***	0.013	0.000
30-Apr-20 – Confirm. Case. Rep.	-0.556***	0.012	0.000
1-May-20 – Confirm. Case. Rep.	-0.491***	0.016	0.000
2-May-20 – Weekend	-0.490***	0.014	0.000
3-May-20 – Weekend	-0.490***	0.015	0.000
4-May-20 – Confirm. Case. Rep.	-0.421***	0.011	0.000
5-May-20 – Confirm. Case. Rep.	-0.396***	0.013	0.000
6-May-20 – JABAR PSBB Announced	-0.395***	0.012	0.000
7-May-20 – Confirm. Case. Rep.	-0.306***	0.018	0.000
8-May-20 – Confirm. Case. Rep.	-0.307***	0.012	0.000
9-May-20 – Weekend	-0.282***	0.014	0.000
10-May-20 – Weekend	-0.258***	0.017	0.000
11-May-20 – Confirm. Case. Rep.	-0.236***	0.012	0.000
12-May-20 – Confirm. Case. Rep.	-0.200***	0.013	0.000
13-May-20 – Confirm. Case. Rep.	-0.169***	0.013	0.000
14-May-20 – Confirm. Case. Rep.	-0.178***	0.013	0.000
15-May-20 – Prior to Ied Mubaraq	-0.136***	0.012	0.000
16-May-20 – Weekend	-0.105***	0.013	0.000
17-May-20 – Weekend	-0.077***	0.015	0.000
18-May-20 – Confirm. Case. Rep.	-0.080***	0.011	0.000
19-May-20 – Confirm. Case. Rep.	-0.060***	0.011	0.000
Observations	25,198		
R-squared	0.988		

*** p<0.01, ** p<0.05, * p<0.1

This regression includes city fixed effects and clustered standard errors at village space and time levels.

Table 2: Social Behavior and Mobility
 Dependent variable: growth rate of active cases

VARIABLES	(1) Coefficient	(2) Std Error	(3) P-value
Lvc	-0.190*	0.115	0.099
lnight_mob	-0.438***	0.043	0.000
d_weekd== 0.0000	0.381***	0.029	0.000
d_holid== 0.0000	0.128**	0.050	0.011
d_bram== 0.0000	-1.042***	0.038	0.000
d_eram== 0.0000	-2.076***	0.035	0.000
Observations	25,198		
R-squared	0.420		

*** p<0.01, ** p<0.05, * p<0.1

This regression includes city fixed effects and clustered standard errors at village space and time levels.

In addition, we need to explore more effects in which particular period the people behave to move from their residence to another area for recreation or culture movement such as holidays (d_holid), weekend (d_weekd), beginning of Ramadhan (d_bram) and end of Ramadhan (d_eram). I present the results in Table 2.

The results will inform us if the mobility is the increase in their resident which means the people are not moving to other regions. The parameter will have a positive sign. Otherwise, if the parameter has a negative sign. The people with higher mobility will move to other regions by reducing the mobility in their original location which higher income levels then move to other regions.

Table 2 indicates the parameter has negative signs of how the rich people move out from their residence to other regions due to culture movement during the beginning of Ramadhan (d_bram) and Before Ied Mubaraq (d_eram). Whereas during weekends and holidays the mobility is increasing in their regions. During this period the confirmed case is increasing significantly. The government has enacted a travel banned and a strong containment policy for reducing human mobility toward outside core COVID-19 areas outside DKI Jakarta and West Java such as Central Java and Eastern Java.

If we looking at the policy news parameter, actually shock news such as travel banned or other policy such as emergency condition was effectively reduced human mobility generally. The news can persuade people to aware of the issues. The government should strongly support health protocol and reducing human mobility as well as the dangers of COVID-19 spread. The transparency of the confirmed case of COVID-19 and death impact should be delivered to the public persistently. The leadership such as the president or higher level of government official such as the Ministry of Health should show up regularly to ensure how government tackle and give awareness for the people to apply health protocol and physical distancing.

Last but not least, is predicting unreported cases. The unreported case is forecasted with two assumptions that are with policy and without policy scenarios. The policy scenario will inform us, whether existing policy such as news, travel banned, and containment policy has reduced unreported case. This unreported case predicted about 1.9 – 2 times of reported case. Especially before Ied Mubaraq, when most people were traveled from Jakarta to West Java, Central Java, Banten, and Eastern Java. The unreported case double significantly.

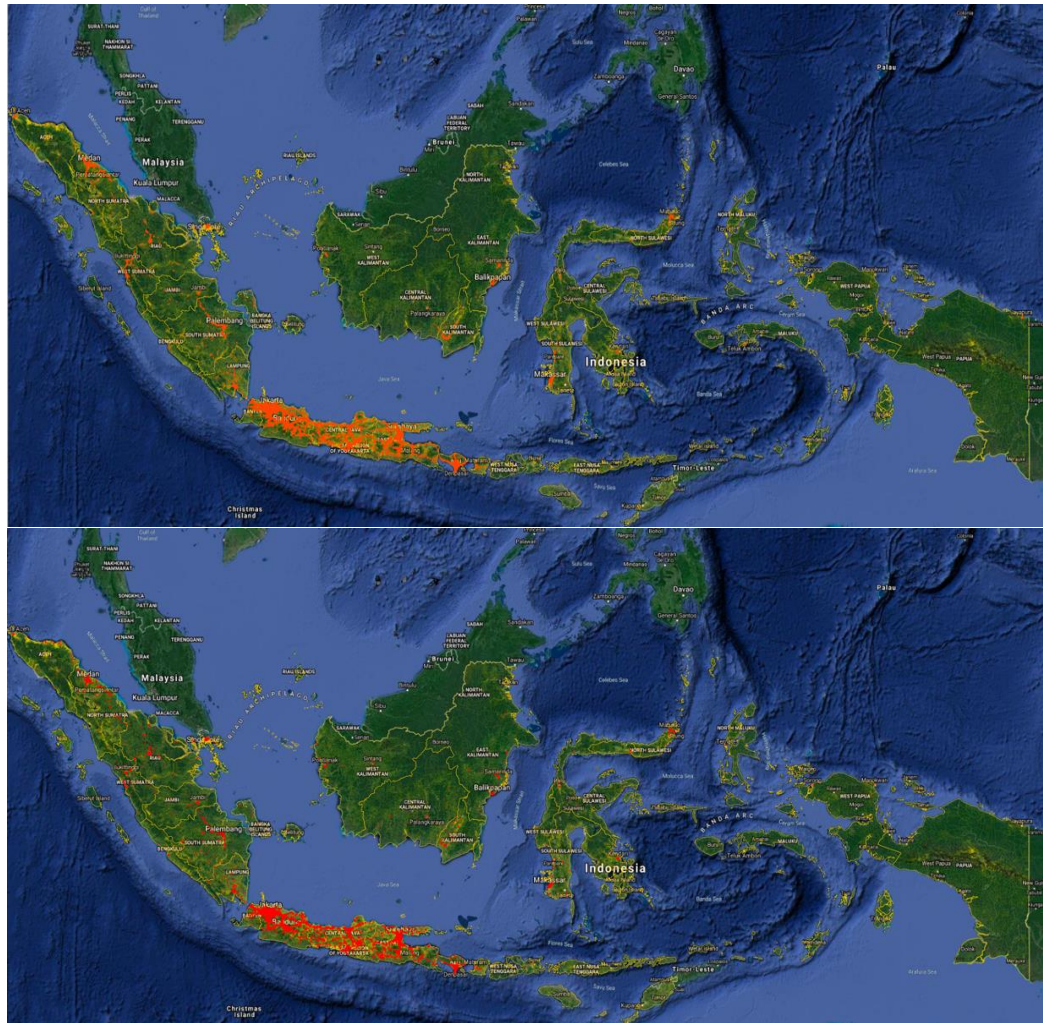


Figure 2: Car Report Congestion Traffic in Indonesia in March-May 2020
Source: Author Estimation with Map Source from Google Map and Waze Mobility (2020)

Without policy action, the unreported case of COVID-19 across Java close to 8,232 million. Whether this number plausible or not, we can compare the results with Hsiang et al., (2020) study. In their study, the number of confirmed cases for China, Iran, and South Korea is extremely large. The confirmed case without containment policy reaches 37 million confirmed cases in China, South Korea about 12 million cases, and Italy about 2.1 million cases.

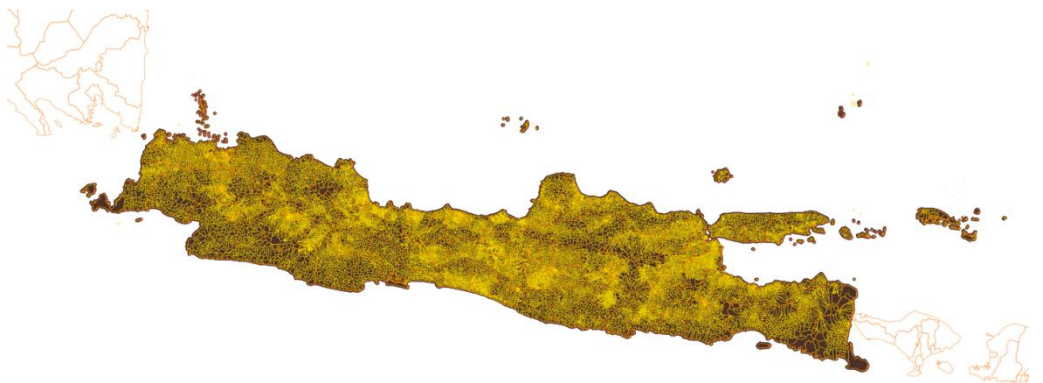


Figure 3: Night Lights Plotting across Village in Java
Source: Author Estimation with Map Source from Google Map and Administration Boundaries from Indonesia Geospatial Information Agency (2020)

Our results support previous studies such as Çolak, Lima, and González (2016) and Charu et al., (2017) and Glaeser et al., (1992) that spatial economic performance is strongly connected to increasing human restrictions mobility. In Indonesia's case, it is confirmed that in the city or higher income has a lack to support the human restriction. The government should strongly support the Health Protocol and the dangers of COVID-19. Indonesian society prefers economics rather than health protocol.

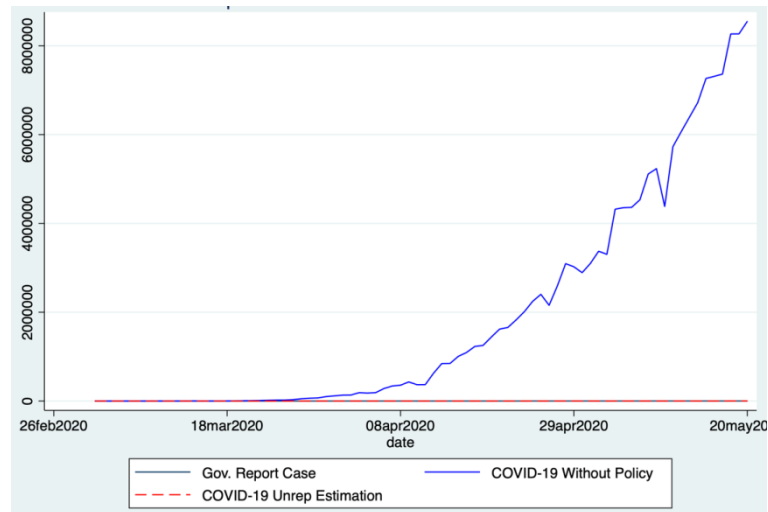


Figure 4: Predicting Unreported Case of COVID-19
Source: Author estimation (2020)

5. Conclusion

We conclude that literacy and awareness of COVID-19 are necessary to be important agenda rather than an economy. The data indicates that Indonesian prefer to engaging with economic rather than health policy. It is very crucial at this moment for Indonesia to more transparent in terms of COVID-19 confirmed case as shock therapy for the society how dangerous this virus. At this moment the research is still based on province level for infected case data whereas mobility data is at the district level. While most of the containment policy at the district level at this moment, it is necessary to improve higher frequency level data for a confirmed case of COVID-19 data. The results could be more accurate and robust to predict how mobility effect on spreading of the COVID-19 case. This paper has a limitation as aforementioned that data frequency in aggregate levels such as provinces. The mobility data is mostly available in Java rather than Outer Java. We need more data collection and other data to confirm with our data such as facebook mobility data.

Acknowledgment

I would like to acknowledge LAPAN for inviting me to publish this paper in KPPA Journal. Furthermore, I would like to appreciate to WAZE Corp that provide me the data as part of the Waze Partner Program.

References

- Charu, V. *et al.* (2017) 'Human mobility and the spatial transmission of influenza in the United States', *PLoS Computational Biology*, 13(2), pp. 1–23. doi: 10.1371/journal.pcbi.1005382.
- Çolak, S., Lima, A. and González, M. C. (2016) 'Understanding congested travel in urban areas', *Nature Communications*, 7. doi: 10.1038/ncomms10793.
- Fields, G. S. (1980). Education and Income Distribution in Developing Countries: A Review of the Literature. *Education and Income: A Background Study for World Development*, 231–315. Retrieved from <http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1487&context=articles>
- Fogli, A. and Veldkamp, L. (2020) 'Germs, Social Networks and Growth', *Review of Economic Studies*, pp. 1–35. doi: 10.2139/ssrn.1679857.
- Gatto, M. *et al.* (2020) 'Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures', *Proceedings of the National Academy of Sciences of the United States of America*, 117(19), pp. 10484–10491. doi: 10.1073/pnas.2004978117.
- Glaeser, E. L. *et al.* (1992) 'Growth in Cities', *Journal of Political Economy*, 100(6), pp. 1126–1152.
- Hanming, F., Wang, L. and Yang, Y. (2020) *Human mobility restriction and the spread of the novel corona virus (2019-NCov) in China*. w26906. Cambridge, MA 02138.
- Henderson, J. V., Storeygard, A. and Weil, D. N. (2012) 'Measuring economic growth from outer space', *American Economic Review*, 102(2), pp. 994–1028. doi: 10.1257/aer.102.2.994.
- Hsiang, S. *et al.* (2020) 'The effect of large-scale anti-contagion policies on the COVID-19 pandemic.', *Nature*. doi: 10.1038/s41586-020-2404-8.
- Lemey, P. *et al.* (2014) 'Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2', *PLoS Pathogens*, 10(2). doi: 10.1371/journal.ppat.1003932.
- Satyakti, Y. (2020) 'Do Human Restriction Mobility Policy in Indonesia Effectively Reduce the Spread of COVID-19', *SSRN Electronic Journal*, (July), pp. 1–30. doi: 10.2139/ssrn.3654553.