Explaining the Effective Reproduction Number of COVID-19 through Mobility and Enterprise Statistics: Evidence from the First Wave in Japan

Yoshio Kajitani[1*] and Michinori Hatayama[2]

[1]Department of Engineering and Design, Kagawa University

[2]Disaster Prevention Research Institute, Kyoto University

*Corresponding author: kajtiani.yoshio@kagawa-u.ac.jp

Abstract

This study uses mobility statistics—a relatively novel data source consisting of smartphone location data—combined with business census data for the eight Japanese prefectures with the highest COVID-19 infection rates to study the effect of lockdown measures on the effective transmission rate of the virus. Based on data for the first wave of infections in Japan, we found that reductions targeting the hospitality industry were more effective than restrictions on general business activities. Specifically, we found that to fully converge the pandemic (that is, to reduce the effective reproduction number to one or less for all the days), a 40-67% reduction in weekly mobility is required, depending on the region. A lesser goal, 80% of days with one or less observed transmission, a 14-61% reduction in weekly mobility is needed.

Introduction

Many countries have suffered from the COVID-19 pandemic and have experienced severe economic impacts due to the restrictions on socio-economic activities. GDP losses have been significant (e.g., -7.9% in Japan during the second quarter of 2020 [1]), and unemployment numbers are increasing. Several countries have managed to restart socio-economic activities close to pre-pandemic levels, but most have suffered a second wave of the pandemic.

The conditions of first, second, and higher-round waves can differ because individual or organizational countermeasures (e.g., masks, hand washing, antiseptic solutions, and partitioning) have advanced. However, analyzing the infection risks and degree of lockdown/voluntary restriction of socio-economic activities in the first wave, the currently available data, is meaningful for creating better activity restriction policies.

In this sense, mobility statistics, which have become recently available through smartphone devices, are a powerful tool for understanding regional overviews of socio-economic activities. For example, Google Mobility Report [2] provides population statistics for retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential areas all over the world. Engle et al. [3] used GPS locational data for 94,116 observations in 3,142 U.S. counties from 2/24/2020 to 3/25/2020 and found that a rise in local infection rate from 0% to 0.003% is associated with a 2.31% reduction in mobility. Yabe et al. [4] employed 200,000 anonymized mobile phone users in Tokyo and concluded that by April 15th (one week into the state of emergency), human mobility behavior decreased by approximately 50%.

Our approach also utilizes mobility data (hourly and 500m grid scale populations all over Japan) considering its powerful ability to capture the conditions of staying home during the pandemic crisis. The focus of the study is two major exploratory data analyses. First, we focus on the question, what types of mobility-restriction measures are correlated with infection risks? For this purpose, recent business census data at a 500m grid scale is combined with mobility statistics and the effective reproduction number (how many people are infected by one infected person, denoted by R(t)). The second focus question is, to what degree do we need to restrict our (daily) travel to reduce the pandemic crisis?

Data set and approach

In this study, three different types of statistics are utilized: the number of infected people [5], mobility statistics [6], and business census data from Japan's Ministry of Internal Affairs and Communications (MIC) [7]. The number of infected people is recorded on the date that the infections are confirmed by Japan's Ministry of Health, Labor and Welfare (MHLW). Here, we focused on the eight prefectures where the number of infections was more than 500 people by May 31, 2020. We then prepared the associated data sets for these eight prefectures.

Fig 1 illustrates the time series of the number of infected people in the eight target prefectures. Tokyo had the most infected people, and the second largest city, Osaka, follows. Explosions of infections can be seen from the end of March. It is assumed that people reduced their restriction levels during the holidays before this large wave came. April 7th is the day the emergency statement was issued by the Japanese Government, after which the first wave of the pandemic gradually abated. The effective reproduction number can be calculated from this data and the serial interval distribution (time between successive infections from one person to another). We employed Cori et al. [8] to estimate the effective reproduction number based on the serial interval distribution provided by Nishiura et al. [9].

Fig 1. Number of Infected People [6].

NTT docomo is one of the largest carriers in Japan, and it holds 37.4 % of all mobile phone contracts in Japan [10]. Populations of people ages 15 to 79 are counted with the following rule for target hourly duration: If a person stays within the grid for 15 minutes during a target hour, then 1/4 is added to the population. Every 15 minutes of stay, this rule is applied. Based on the duration of each person's stay in a grid, either 1/4, 1/2, 3/4, or 1 is added to the population of the grid. The mobility data are not raw data, but are magnified based on the carrier's, NTT docomo, share in each region.

Fig 2 depicts the two different timings for the mobility statistics in Tokyo: from 15:00-16:00 on March 13 (Friday), a relatively dense population is observed, while in the same time period on Apr 23 (Friday), the population became smaller in the center of Tokyo. The total number of people observed in the grids during the time period in March is 12,943,780 people[1], while 12,128,864 people were counted in April. Reductions are seen in inflows from other prefectures and abroad as well as in inflows from residential areas to the center of the prefecture.

Fig 2-1. Mobility statistics in Tokyo at 15:00 (Mar 13, 2020).

Fig 2-2. Mobility statistics in Tokyo at 15:00 (Apr 23, 2020)

The last statistic introduced is the 2016 economic census for businesses by the MIC [7]. The statistics include the number of employees in over 100 business sectors, and it is aggregated in a 500m grid scale. By using the mobility statistics and business census, we use the following criterion as a measure of potential contacts in the business and commercial districts. This criterion should have a strong (negative) correlation to the level of stay home activity. The measure of potential contacts at the business and commercial districts is defined as

$$\mathrm{PC}(\mathrm{S},\mathrm{T}) = \frac{\sum_{t \in T} \sum_{s \in S} Emp(s)\big(N(s,t)\big)^{\alpha}}{\sum_{s \in S} Emp(s)}, \tag{1}$$

where $Emp(s)$ the number of employees at grid s, and $N(s,t)$ is the mobility statistics at grid s and

[1] A 2015 national census survey indicates a daytime population in Tokyo of 15,920,405 (Tokyo metropolitan government, 2018) [11].

time t. If people are crowded in business and commercial areas, where the number of employees is large, the value of the measure becomes large. In the analysis, we investigate two cases: $Emp(s)$ as the total employees in all business sectors and as the total employees in all hospitality sectors (wholesale and retail, hotel and restaurant, living related and personal services, amusement, education, and medical and healthcare sectors).

Results and discussion

First, the value of the parameter for Equation (1) is determined by maximizing the correlation between the PC index and R(t) for two cases of $Emp(s)$. R(t) is estimated as a weekly average before t (*i.e.,* R(t) represents the number from day t-6 to t). Because the estimates of R(t) in the first few weeks have a wide confidence interval due to the small number of incidences, we set the target period to the days from March 15 to May 31, 2020.

Fig 3 describes the Peasron correlation coefficients between PC and R(t) at different parameter scales α. PC is also estimated as a weekly average because it shows generally high correlation with R(t). The correlation indices are estimated for each prefecture first and then averaged among the eight prefectures. From this figure, the correlation coefficients are generally better when employees in the hospitality sector are selected as the weight of PC. This indicates that infections tend to occur in the hospitality sector. In the case in which the hospitality sector is used for Equation (1), α=8 performs the best. The reason behind this needs to be investigated further, but the parameter values may represent the frequency and time duration of contacts among people in a grid.

Fig 3. Pearson correlation coefficients between effective reproduction numbers (R(t)) and the measure of potential contacts in the business and commercial district at different power scales, α in Equation (1) (applied to the daily data set from Mar 15 to May 31, 2020).

Fig 4 plots the daily time series of the measure of potential contacts and R(t) in the eight prefectures. In the figure for Hokkaido, the target days after March 15th are highlighted for a reference. In many prefectures, the potential contacts decrease considerably during April and the beginning of May, but gradually start to return to pre-pandemic conditions at the end of May. The state of emergency declared by the central government ended on May 16th in Hokkaido and on May 31st in all the other prefectures, but people gradually restarted their activities probably because the atmosphere of emergency was alleviated after many of the other prefectures ended emergency actions.

(Fig 4-1.

Fig 4-2.

Fig 4-3.

Fig 4-4.

Fig 4-5.

Fig 4-6.

Fig 4-7.

Fig 4-8.)

Fig 4. Effective reproduction number and measure of potential contact in the eight prefectures

Based on the estimates of PC and R(t) described in Fig 4, the required mobility restriction levels can be calculated for converging the pandemic. That is, the threshold value of PC can achieve R(t) ≤ 1. If the threshold value of PC is set as PC', then $PC'^{\frac{1}{\alpha}}/PC^{\frac{1}{\alpha}}$ gives the relative ratio of the weighted geometric mean of N'(s,t) and N(s,t), where N'(s,t) is the required population at grid *s* at time *t* and one of the solutions to achieve the threshold value, assuming N(s,t) is population at normal period. Our study adopts average N(s,t) in February, 2020. Fig 5 shows the estimated population (mobility) restriction levels (N'(s,t)/N(s,t)) to ensure achievement of R(t) ≤ 1 in each prefecture. Because this is the geometric mean scale weighted by employees in the hospitality sector, PC' can be more easily achieved if the rate of people is reduced more in places where the hospitality sector is agglomerated.

From the figure, Osaka requires the largest reduction (67%), followed by Tokyo. This result can be naturally interpreted as these prefectures are the largest in Japan, and the population in the hospitality sector tends to be large. As shown in Fig 4, the scale of potential contacts in these prefectures is more than 10 times larger than the scale in other prefectures. Kanagawa, Hyogo, and Fukuoka also require high restriction levels on visits to the hospitality sector. Among these prefectures, Hyogo is a less populated prefecture, and the index of PC is low. Population characteristics are generally reflected in the low value of R(t) in Hyogo, but a large restriction on visits to the hospitality sector is required to guarantee R(t) ≤ 1. Another index, such as R(t) ≤ 1 with an 80% chance, may be appropriate to capture the relationships between PC and an average low value of R(t). Hokkaido, Saitama, and Chiba may be classified into the third group, where the required restrictions are not so strict. In these prefectures, Chiba maintained relatively high values of R(t) at the end of March, but effectively reduced it through small population reductions in hospitality sector areas.

Fig 5. Necessary reduction level of visits to hospitality sectors to achieve R(t)<1 for "100% of days"

Based on the discussion above, in order to better understand the case of Hyogo, another index "R(t) ≤ 1 for 80% of days" is introduced to capture the relationships between PC and an average low value of R(t) in Fig 6. This provides a slightly different view from Fig 5. The required reduction level becomes much lower, especially in Saitama, Chiba, Osaka, Hyogo, and Fukuoka, where generally low R(t)s are observed. On the other hand, Hokkaido, Tokyo, and Kanagawa still require a large reduction in number of visits to the hospitality sector. In these prefectures, a large number of infections were observed from an early stage, and the introduction of countermeasures (e.g., the number of people wearing masks and social distancing at local spots within a 500m grid) might have been preventative.

Fig 6. Necessary reduction level of visits to hospitality sectors to achieve R(t)<1 for "80% of days"

The above result is based on the case in which an 80% reliability level is arbitrarily determined, but more discussions may need to determine the reliability level that we can accept to determine the mobility reduction. A better discussion would be to determine the relationships between the mobility reduction levels at different reliability levels for R(t)<1 and estimate the economic impacts of different policies. However, this is beyond the scope of this study.

Another necessary discussion point lies in the difference between the first wave treated in this study and the second wave (after the end of June). Considering that the countermeasures have advanced, less restrictions on mobility may achieve R(t)<1. Continuous monitoring is necessary to understand when we will establish a new life with COVID-19.

Conclusions

This study utilized mobility statistics and a business frame census on a fine spatial scale to capture the effective reproduction number of COVID-19, which is an important indicator in epidemiology. The measure of potential contacts in the hospitality sector/total business sector was defined, and the values of its parameters were estimated by maximizing the correlation between the measures and the effective reproduction number in eight Japanese prefectures, where the incidence is large. One of the major conclusions in this study is that the measure of potential contacts in the hospitality sector has a fair correlation with the effective reproduction number.

From this measure, the necessary population reduction level to converge the pandemic can be derived. Our analysis indicated 0.40-0.67 reductions are required to achieve R(t)<1 for all the days, depending on the conditions of the prefectures, but 0.14-0.61 are enough to achieve R(t)<1 for 80% of the days. Because of the regional variety in values, and high sensitivity to the required reliability to achieve R(t)<1, these relationships should be carefully checked in each prefecture to determine mobility restriction policies. An analysis of the relationships between mobility reduction and economic impacts would also assist in this kind of policy making.

For additional future studies on Japanese conditions, a comparative study between the first and second waves will be important to identify the progress of countermeasures. A similar analysis in other countries would also help to understand what level of mobility restrictions and local countermeasures would contribute to a low infection risk.

Acknowledgements

We are grateful to NTT docomo InsightMarketing Inc. for providing mobility data sets in this study.

References

1. Cabinet Office, Japan. Quarterly estimates of GDP. https://www.esri.cao.go.jp/en/sna/sokuhou/sokuhou_top.html, Accessed 16 September 2020.
2. Google. Community mobility reports, https://www.google.com/covid19/mobility/, Accessed 13 September 2020.
3. Yabe T, Tsubonouchi K, Fujiwara N, Wada T, Sekimoto, Y, and Ukkusuri, SV. Non-compulsory measures sufficiently reduced Human mobility in Tokyo during the Covid-19 epidemic, arXiv:2005.09423, https://arxiv.org/abs/2005.09423, Accessed 13 September 2020.
4. Engle S and Stromme J and Zhou A Staying at home: Mobility effects of COVID-19 (April 3, 2020). http://dx.doi.org/10.2139/ssrn.3565703
5. NTT docomo, Mobile spatial statistics, https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/, Accessed 13 September 2020.
6. Ministry of Health, Labor and Welfare (MHLW): About Coronavirus Disease 2019 (COVID-19), https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/newpage_00032.html, Accessed 13 September 2020.
7. Ministry of Internal Affairs and Communications (MIC). 2014 Economic Census for Business Frame. http://www.stat.go.jp/english/data/e-census/index.htm, Accessed 13 September 2020. (Data Provider: Statistical Information Institute for Consulting and Analysis).
8. Cori A, Ferguson NM, Fraser C, and Cauchemes S. A new framework and software to estimate time-varying reproduction numbers during epidemics, American Journal of Epidemiology. 2013; 178(9): 1505-1512.
9. Nishiura H, Linton NM and Akhmetzhanov AR. Serial interval of novel coronavirus (2019-nCoV) infections, medRxiv 2020, https://www.medrxiv.org/content/10.1101/2020.02.03.20019497v2, 13p.

10. Ministry of Internal Affairs and Communications (MIC). Contract shares in communication service. https://www.soumu.go.jp/menu_news/s-news/01kiban04_02000162.html, Accessed 13 September 2020.
11. Tokyo Metropolitan government (2018): Daytime population at Tokyo, https://www.metro.tokyo.lg.jp/tosei/hodohappyo/press/2018/03/20/12.html, Accessed 13 September 2020.

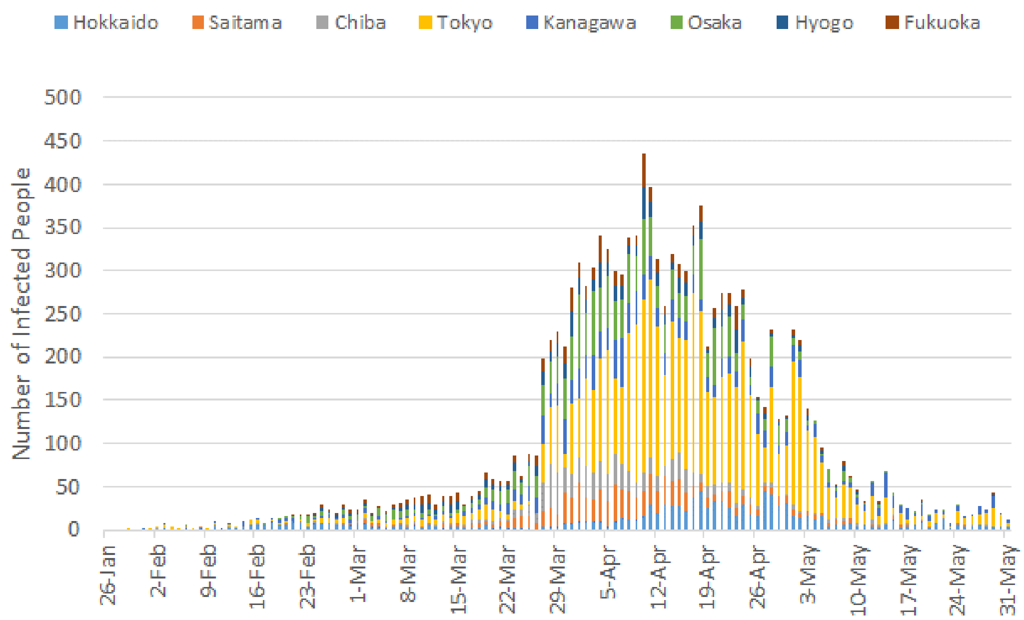


Fig 1. Number of Infected People [6]

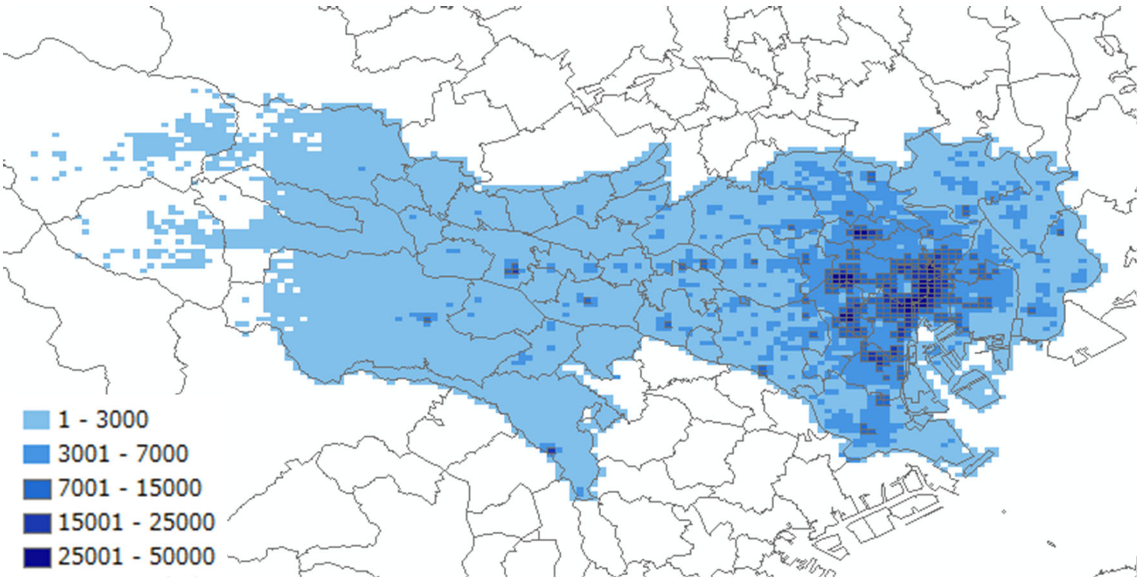


Fig 2-1. Mobility statistics in Tokyo at 15:00 (Mar 13)

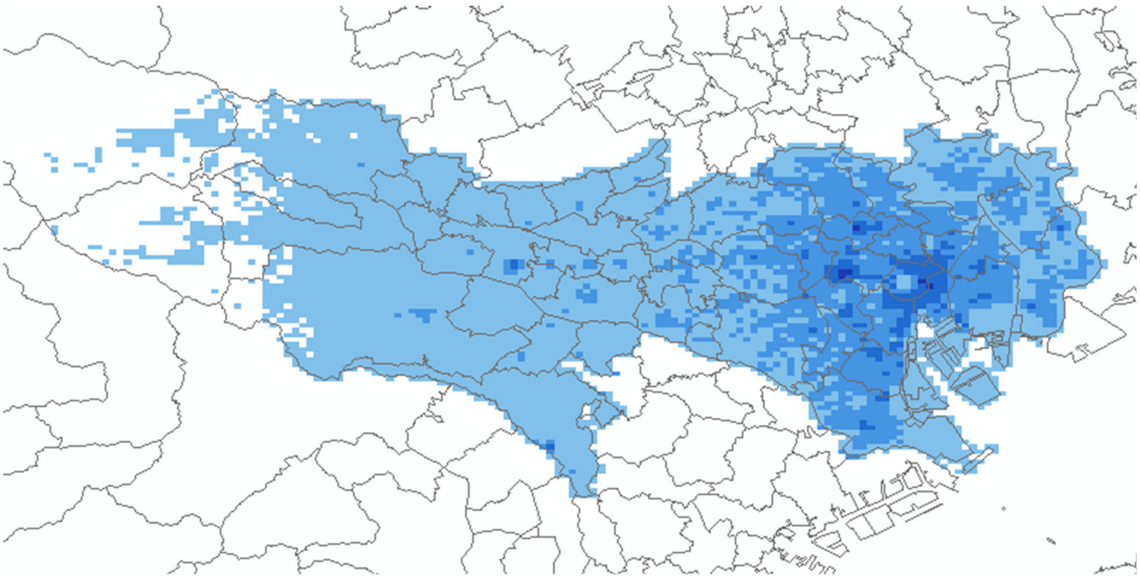


Fig 2-2. Mobility statistics in Tokyo at 15:00 (Apr 23)

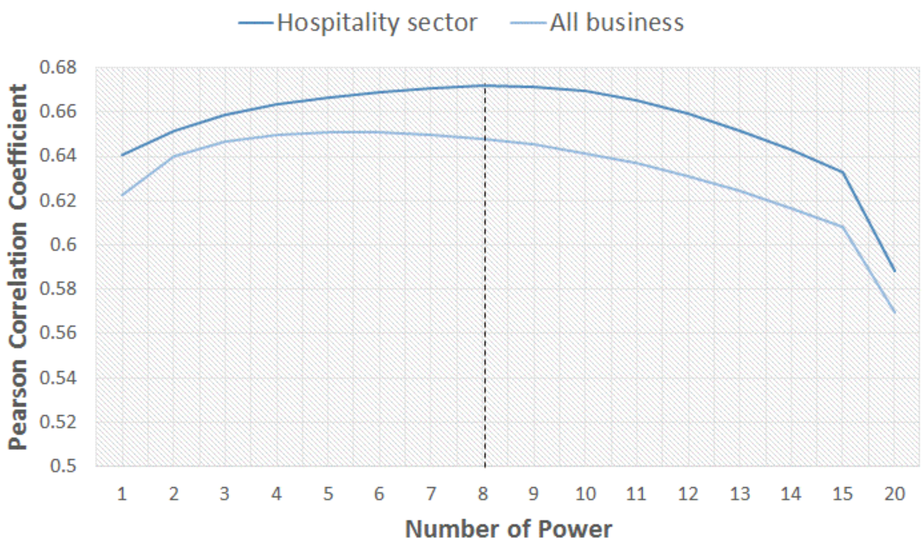
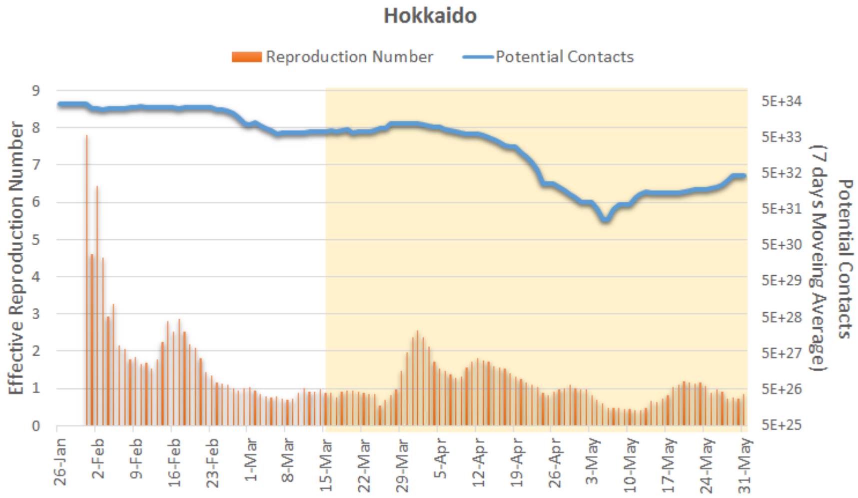
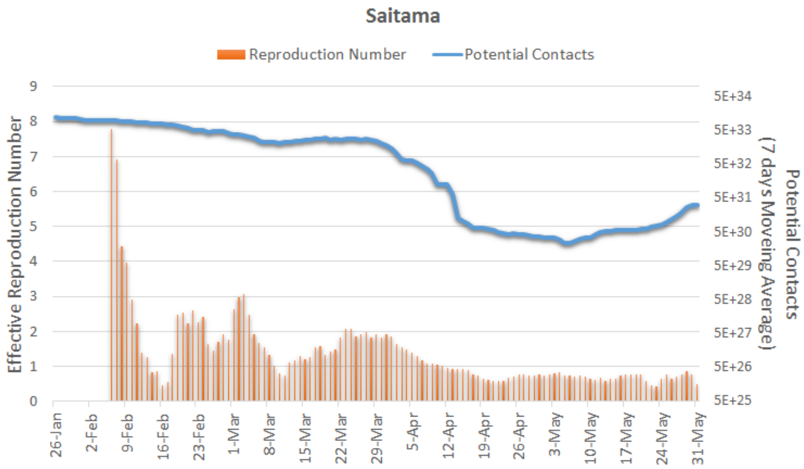
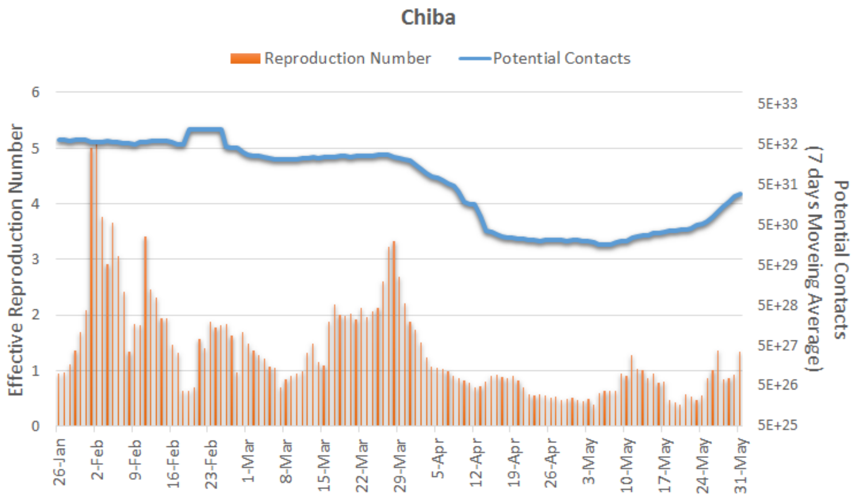
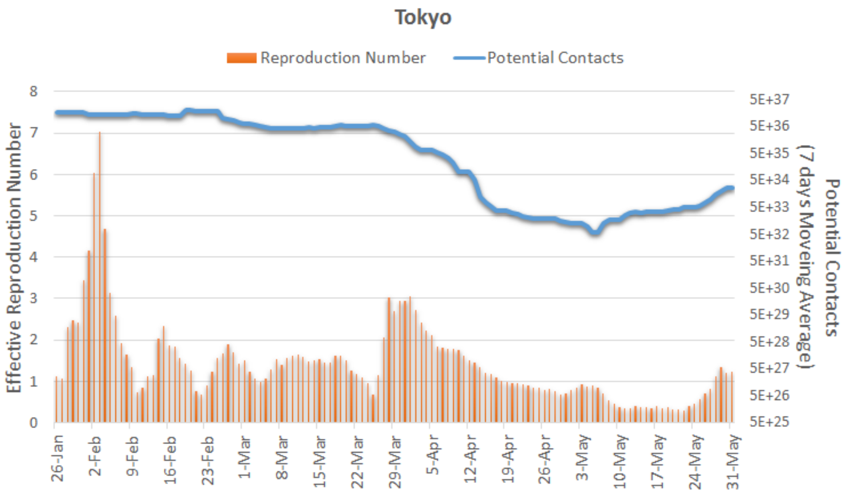


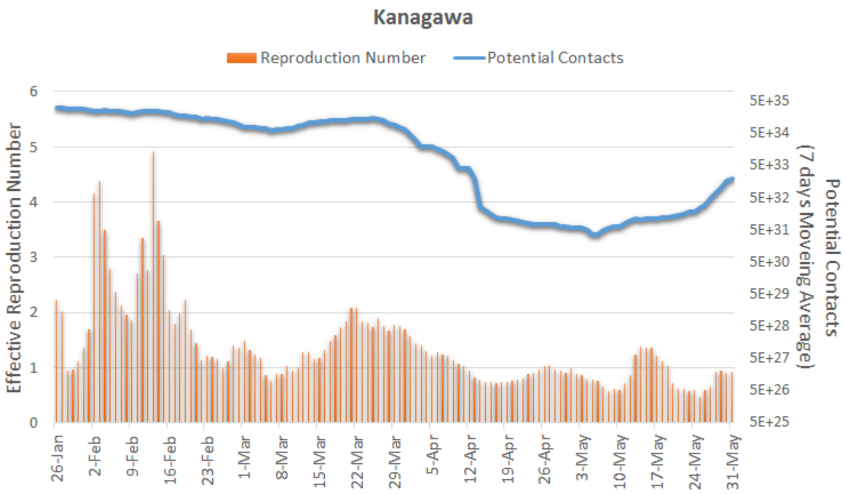
Fig 3. Pearson correlation coefficients between effective reproduction numbers (R(*t*)) and the measure of potential contacts in the business and commercial district at different power scales, α in Equation (1) (applied to the daily data set from Mar 15 to May 31, 2020).

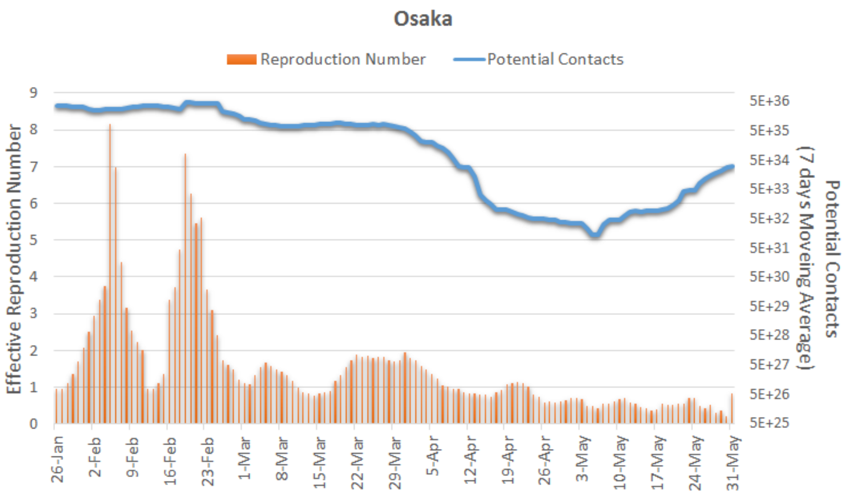

Hokkaido
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts
(7 days Moveing Average)

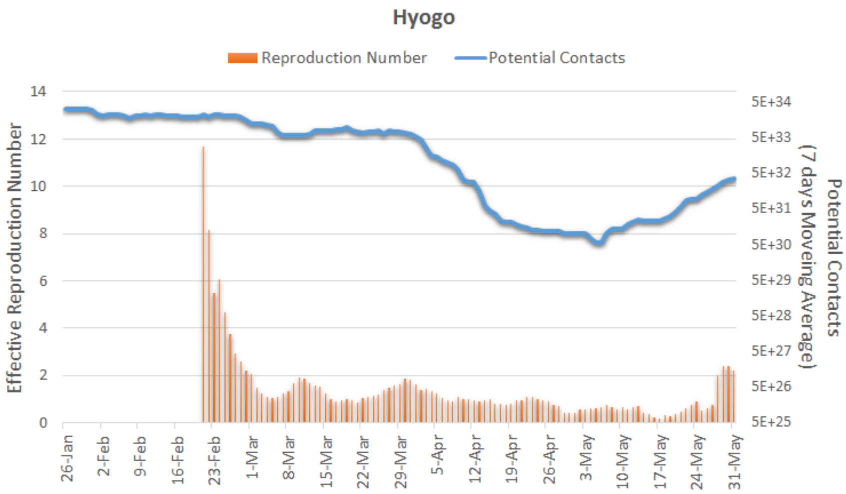

Saitama
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts
(7 days Moveing Average)

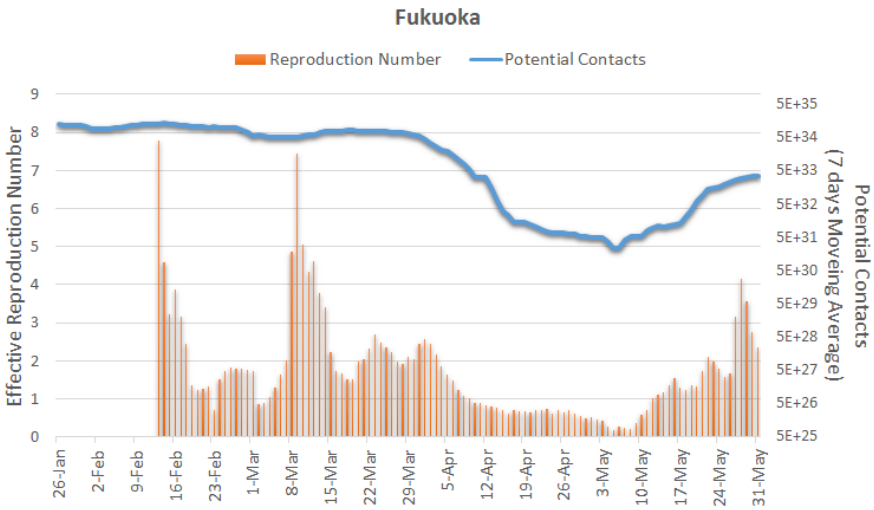

Chiba
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts
(7 days Moveing Average)


Tokyo
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts (7 days Moveing Average)
0
1
2
3
4
5
6
7
8
5E+25
5E+26
5E+27
5E+28
5E+29
5E+30
5E+31
5E+32
5E+33
5E+34
5E+35
5E+36
5E+37
26-Jan
2-Feb
9-Feb
16-Feb
23-Feb
1-Mar
8-Mar
15-Mar
22-Mar
29-Mar
5-Apr
12-Apr
19-Apr
26-Apr
3-May
10-May
17-May
24-May
31-May


Kanagawa
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts (7 days Moveing Average)
0
1
2
3
4
5
6
5E+25
5E+26
5E+27
5E+28
5E+29
5E+30
5E+31
5E+32
5E+33
5E+34
5E+35
26-Jan
2-Feb
9-Feb
16-Feb
23-Feb
1-Mar
8-Mar
15-Mar
22-Mar
29-Mar
5-Apr
12-Apr
19-Apr
26-Apr
3-May
10-May
17-May
24-May
31-May


Osaka
Reproduction Number
Potential Contacts
Effective Reproduction Number
Potential Contacts (7 days Moveing Average)
0
1
2
3
4
5
6
7
8
9
5E+25
5E+26
5E+27
5E+28
5E+29
5E+30
5E+31
5E+32
5E+33
5E+34
5E+35
5E+36
26-Jan
2-Feb
9-Feb
16-Feb
23-Feb
1-Mar
8-Mar
15-Mar
22-Mar
29-Mar
5-Apr
12-Apr
19-Apr
26-Apr
3-May
10-May
17-May
24-May
31-May





Figs 4-1-4-8. Effective reproduction number and measure of potential contact in the eight prefectures

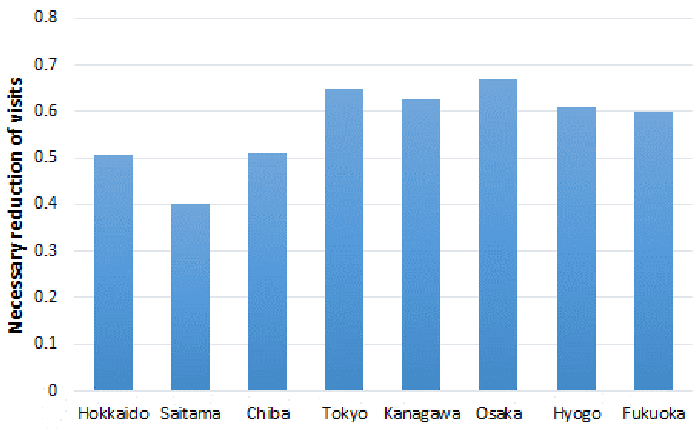


Fig 5. Necessary reduction level of visits to hospitality sectors to achieve R(t)<1 for "100% of days"

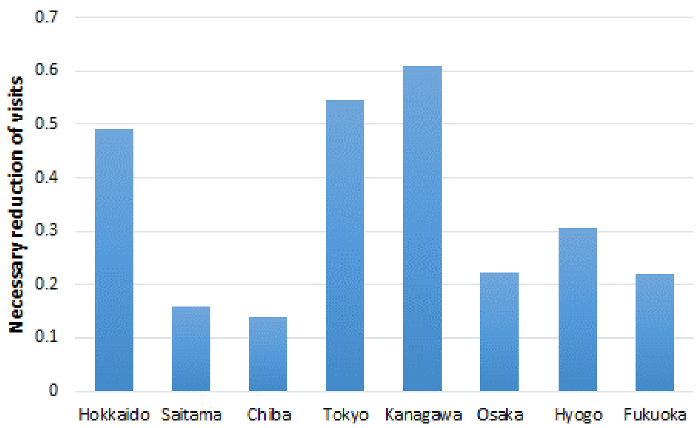


Fig 6. Necessary reduction level of visits to hospitality sectors to achieve R(t)<1 for "80% of days"