

The impact of mass gatherings on the local transmission of COVID-19 and the implications for social distancing policies: Evidence from Hong Kong

Abstract

Mass gatherings provide conditions for the transmission of infectious diseases and pose complex challenges to public health. Faced with the COVID-19 pandemic, governments and health experts called for suspension of gatherings in order to reduce social contact via which virus is transmitted. However, few studies have investigated the contribution of mass gatherings to COVID-19 transmission in local communities. In Hong Kong, the coincidence of the relaxation of group gathering restrictions with demonstrations against the National Security Law in mid-2020 raised concerns about the safety of mass gatherings under the pandemic. Therefore, this study examines the impacts of mass gatherings on the local transmission of COVID-19 and evaluates the importance of social distancing policies. With an aggregated dataset of epidemiological, city-level meteorological and socioeconomic data, a Synthetic Control Method (SCM) is used for constructing a ‘synthetic Hong Kong’ from over 200 Chinese cities. This counterfactual control unit is used to simulate COVID-19 infection patterns (i.e., the number of total cases and daily new cases) in the absence of mass gatherings. Comparing the hypothetical trends and the actual ones, our results indicate that the infection rate observed in Hong Kong is substantially higher than that in the counterfactual control unit (2.63% vs. 0.07%). As estimated, mass gatherings increased the number of new infections by 62 cases (or 87.58% of total new cases) over the 10-day period and by 737 cases (or 97.23 %) over the 30-day period. These findings suggest the necessity of tightening social distancing policies, especially the prohibition on group gathering regulation (POGGR), to prevent and control COVID-19 outbreaks.

Keywords: COVID-19 pandemic; Mass gatherings; Social distancing policy; Prohibitions on group gathering

Introduction

Managing the complex and unique health security risks posed by mass gatherings is challenging (Blyth et al., 2010; Ishola & Phin, 2011; Rainey et al., 2016). Many studies point out that mass gatherings can be associated with increased risks and amplified transmission of infectious diseases which are transmitted through respiratory routes, such as influenza, measles, and meningitis (Al-Tawfiq & Memish, 2012; Memish et al., 2015; Yezli et al., 2016; Hong & Gautret, 2018; Gautret, 2020). Under the circumstance of the COVID-19 pandemic, mass gatherings, according to the World Health Organization (WHO), amplify the spread of viruses and inhibit a country's ability to respond (WHO, 2020). It is well established that COVID-19 is transmitted between people through direct/indirect contacts, contaminated subjects and environmental factors, and the infection risks are proportional to the closeness of interactions with people who are infected (Li et al., 2020; Ong et al., 2020; Burke et al., 2020; Li et al., 2021). A stricter risk assessment recommended by WHO states that even medium-sized gatherings may create conditions that enable the transmission of COVID-19 (WHO, 2020). Given the high contagiousness of COVID-19, crowded gatherings can create environments conducive to virus transmission among participants and lead to subsequent dissemination within their families. When a community is experiencing an outbreak, mass gatherings may lead to a reduction in social distancing behavior and create challenges for the prevention and control of infectious diseases. Notwithstanding subsiding pandemic situations, public health officials still warn that mass gatherings, with extensive social mixing and uncertainties, has the potential to trigger a resurgence of COVID-19 infections. The risks of infections through mass gatherings are also partly determined by the type, venue, and location of mass gatherings and the demographics of participants (Shi et al., 2010; Abubakar et al., 2012; Ebrahim & Memish, 2020). The gatherings that involve closer and more extended contact and are held indoors or at overcrowded sites, such as professional conferences and musical events, may be more susceptible to infections. Other types of gatherings such as demonstrations and sporting events, which often take place in outdoor environments with better ventilation, may have lower transmission risks.

In 2020, social distancing and the exhortation to avoid social contact were adopted worldwide to contain the transmission of COVID-19. Many jurisdictions further imposed restrictions or bans on group gatherings as part of their anti-pandemic policy. In Hong Kong, the government imposed the Prohibition On Group Gathering Regulation¹ (POGGR) from March 2020 onwards, stipulating the number of people allowed in a group and banning large-scale mass gatherings. Case clusters of COVID-19 were reported worldwide in parties and personal or social activities where a relatively small number of people were involved. To date, the scientific community has gained a significant understanding of how these small group gatherings in indoor settings could contribute to the virus transmission (see for example, Furuse et al., 2020; Lelieved et al., 2020; Vuorinen et al., 2020; Moritz et al., 2021). But to what extent large-scale gatherings may contribute to the spread of COVID-19 in local

¹ Associated regulations are made by the Chief Executive in Council under section 8 of the Prevention and Control of Disease Ordinance (Cap. 599).

communities? Although it is widely believed that mass gatherings could disrupt social distancing and exacerbate the scope of the virus transmission due to the complex patterns of social mixing at these events (Ishola & Phin, 2011; Rainey et al., 2016; Ebrahim & Memish, 2020), empirical evidence to date is rather limited. Mat et al. (2020) analyzed a single mass gathering, the four-day Sri Petaling Mosque religious event in Malaysia, and concluded that it was a key catalyst for the second wave of COVID-19 outbreak in the country. A retrospective cohort study by Domènech-Montoliu et al. (2021) also found that the COVID-19 transmission and subsequent outbreak in Borriana, a municipality in Spain, were associated with several mass gathering events related to a traditional festival. In contrast, opposite evidence and conclusions also exist. A recent study suggests that the *Black Lives Matter Public Gatherings* in the U.S. increased non-participants' stay-at-home behavior such that it offset the increased transmission risks among protestors (Dave et al., 2020). In summary, there is still a dearth of empirical research for performing a comprehensive assessment of the impact of mass gatherings on the transmission of COVID-19 in local communities. The purpose of this study is to provide additional empirical evidence and reduce this gap in the literature.

In this study, we pay special attention to public gatherings that occurred in Hong Kong in mid-2020 when public tension escalated due to the enactment of the National Security Law (NSL). During the COVID-19 outbreak, the coincidence of the relaxation of POGGR and public demonstrations against the NSL in Hong Kong provides us with an opportunity to study the impact of mass gatherings on the local transmission of COVID-19. Using Hong Kong as a case study, this study examines the epidemiological data to probe any correlation between mass gatherings and increases in COVID-19 local cases.

Policy environment and public gatherings in Hong Kong

Hong Kong has a unique political status and subtle relationship with mainland China under the "One Country, Two Systems" principle, which safeguards national sovereignty while allowing Hong Kong a great degree of autonomy, retaining its own economic and administrative systems. In Hong Kong, public demonstrations are held yearly on July 1, usually organized by the pro-democracy camps (Cantoni et al., 2016; 2019). In 2019, Hong Kong experienced the most severe bout of social unrest in the city's history. The *Anti-Extradition Law Amendment Bill Movement* proceeded in the form of public gatherings and demonstrations and was eventually wiped out in the same year because of a government crackdown. However, political tension again mounted between June 2020 and July 2020, after Beijing's decision to promulgate the NSL.

Faced with the pandemic, the Hong Kong government has implemented various social distancing policies to control community transmission. The most significant is the Prohibition On Group Gathering Regulation (POGGR), first implemented on March 29, 2020, which

forbids any form of gathering of more than four people in any public place². With Hong Kong experiencing social unrest over the past two years, social distancing policies not only reduced regular social gatherings such as dining but, more importantly, prohibited and restricted public gatherings and demonstrations. When the outbreak stabilized from mid-April through late June of 2020, the Hong Kong government actively reviewed and revised its social distancing policies. Under the "Suppress and Lift" strategy, the government reviewed the feasibility of relevant measures and made amendments to restrictions on mass gatherings on a timely basis. Infection risks of certain activities and premises/places, and overseas practices were also taken into account in such decision making. As the confirmed case numbers were on a decline at the time, the government decided to relax restrictions. The maximum allowable size of group gatherings under the POGGR was raised to 8 people on May 8, and then further lifted to 50 people on June 19. Modifications in these social distancing policies are reflected in the frequency and scale of mass gatherings held during the period. In fact, in relation to the relaxation of social distancing orders, different categories of population movement including transport, parks, and shopping all showed an upward trend in July 2020, according to the Google Mobility data³.

Concurrent with the relaxation of social-distancing restrictions (e.g., the maximum size for group gatherings raised to 50 people) and the enactment of the NSL, massive demonstration events reappeared in the city in late June and early July of 2020. Beginning on June 20, public gatherings and demonstrations against the NSL were frequently held in different areas of the city. Figure 1 summarizes the major gatherings that occurred during this period. The largest anti-NSL march took place on July 1, the date when anti-authoritarian public gatherings are held every year in the past. According to local media, an estimate about tens of thousands of protesters came onto the streets of Hong Kong Island to protest against the enactment of the NSL. The government reported that around 370 people were arrested by police officers.⁴ Although participants in these demonstrations and protest activities usually adhered to certain precautions such as wearing masks, strictly maintaining social distancing norms was difficult since these gatherings were subject to unpredictable turns of events, including confrontation and violence. While the pandemic was still ongoing and the government's *Stay-at-home Initiative* was still active, Hong Kong's opposition camp organized an unofficial poll, the proclaimed pro-democratic primary election, on July 11 and 12, convening more than 20,000 people to join the offline voting.⁵ In a nutshell, the conjuncture of the lift of POGGR and the announcement of the NSL collectively triggered the emergence of public gatherings and demonstrations in Hong Kong, which could have created a risk of virus transmission. Therefore, in this paper, we try to uncover the impact of relaxing social distancing policies and the subsequent mass gathering events on COVID-19

² Public place is defined in the Regulation as "a place to which the public or a section of the public may or are permitted to have access from time to time, whether by payment or otherwise".

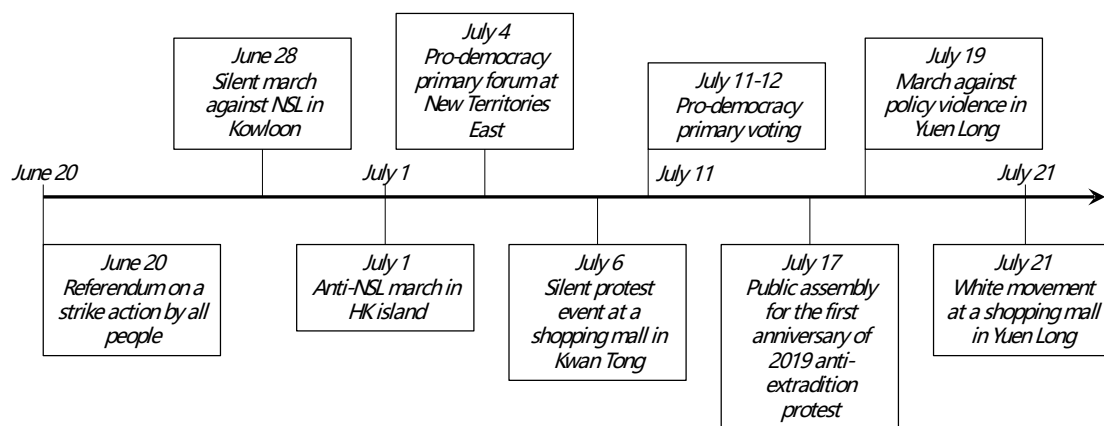
³ For inside-city population movement data provided by the Google, see <https://www.google.com/covid19/mobility/>

⁴ HKSAR Government Press Release, <https://www.info.gov.hk/gia/general/202007/01/P2020070200001.htm>

⁵ The number of valid ballots recorded throughout the two-day vote was provide by the opposition camp (see <https://www.nbcnews.com/news/world/over-600-000-hong-kong-cast-protest-vote-against-new-n1233614>). The number was not recognized by any official governmental departments.

transmission in Hong Kong. It should be noted that other forms of gatherings and social interactions had also resumed to a certain degree following the relaxation of social distancing orders back then. However, mass gatherings are characterized by greater uncertainty and risks of viral spread because of high degree of social mixing. In fact, many cases in the third wave are recorded as infection from unknown sources (comprise as high as almost 40% of total infections).⁶ There are certainly invisible transmission chains and environmental factors that contributed to the large scale of transmission in the third wave in Hong Kong. According to the Health Bureau, the third wave of outbreak mainly stems from community infections and covers a wide range of regions and industries.⁷ Therefore, the research objective of this study is to investigate the overall impact of mass gatherings on the third wave of COVID-19 local outbreaks in Hong Kong after the lift of social distancing policy (i.e., the POGGR).

Fig 1: The timeline of a series of mass gatherings and public demonstrations in Hong Kong in mid-2020

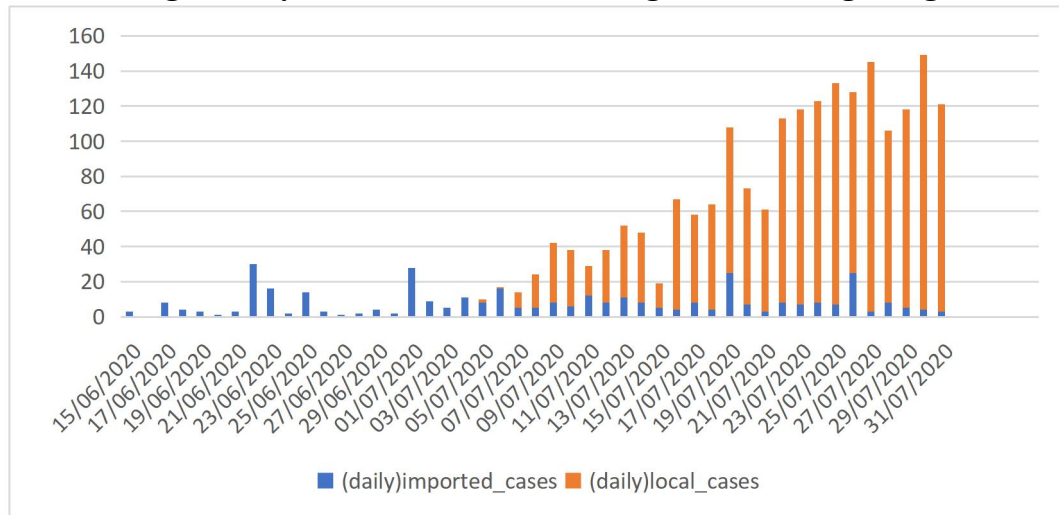


It is also noted that Hong Kong has implemented restrictions on international and cross-border travel and imposed compulsory quarantine requirements on people coming into the city to manage the risk of case importation. Given that imported cases were largely blocked, many surmise that the rebound and catastrophic proliferation in local cases (see Fig 2) may be related to these mass gathering events and public demonstrations.

⁶ Chief Executive Carrie Lam: Active participation in testing urged
https://www.news.gov.hk/eng/2020/08/20200825/20200825_122029_411.html

⁷ Health Bureau. https://www.healthbureau.gov.hk/blog/en/2020/post_20200823.html

Fig 2: Daily COVID-19 new cases diagnosed in Hong Kong



Source: Centre for Health Protection under the Health Department,
<https://www.chp.gov.hk/en/features/102997.html>

Epidemiological trajectory of COVID-19

From mid-April to late June of 2020, the outbreak in Hong Kong had stabilized, but beginning July 5, the outbreak began to re-erupt, with an upsurge in COVID-19 infections (mostly classified as local cases). The situation suddenly deteriorated into the largest outbreak Hong Kong had then seen, which was commonly referred to as "*the third wave*" of local outbreak. The figure of daily new infections reached another peak on July 27 (145 new cases) and the city's daily new case numbers remained in the triple digits for 7 consecutive days, bringing the city's total number of infections to 2884 and total related deaths to 23 as of July 28. Moreover, most of the infections reported in July were local cases, and the source of infection for more than 40% of these cases was unknown (Gov. HK, 2020). Figure 2 shows the changes in new confirmed cases of COVID-19 in a daily basis in Hong Kong from June 15 to July 31, 2020. It is unclear why the pandemic in Hong Kong suddenly rebounded after the city's initial success in controlling the outbreak. While mainland China has been able to keep the spread of COVID-19 under control since May 2020, Hong Kong and places such as the United States and Europe have seen a second or third wave of the outbreak. Yet, we are still unclear about the source and transmission routes of the third wave of the local outbreak in Hong Kong.

Given all these, this study investigates the question of whether mass gatherings, particularly in the form of public gatherings and demonstrations, contributed to the third wave of local outbreak of COVID-19 in Hong Kong. If our results suggest that mass gatherings have contributed to the increase in COVID-19 cases, it would be reasonable to strengthen social distancing measures, for example, group gathering bans, as a key strategy in containing the COVID-19 outbreak. In this regard, our study can generate useful insights on pandemic control for countries considering or starting to relax their restrictions on mass gatherings to adapt to the new normal amid this pandemic. Furthermore, our findings may help

governments improve their policymaking capacity to combat other novel infectious diseases in the future. Large-scale public demonstrations in Hong Kong during mid-2020 and the associated consequences can be a reference for policymaking in many places to manage various mass gatherings such as religious events in the face of other new epidemics.

Method and Materials

We utilize a quantitative approach, namely, Synthetic Control Method (SCM) for a comparative study to assess the impact of mass gathering on the local spread of COVID-19 in the local communities of Hong Kong. This analysis has a premise that the only way we can assess the influence of mass gatherings on the local transmission of COVID-19 is to compare the policy outcome with an appropriate counterfactual control unit (Adhikari et al., 2018). In reality, it would be difficult to find a city sufficiently similar to Hong Kong to act as the control unit. By selecting a combination of possible comparison units that is most representative of Hong Kong's characteristics before the interventions, SCM suggests a systematic way of reproducing the counterfactual (Zhu & Tan, 2021, 2022). To achieve this, an epidemiological database and an aggregator of city-level demographic, socioeconomic, and climatic data for more than 200 Chinese cities (including cities in the mainland and Macau) were crosswalked. Based on this combined dataset, an SCM was then employed to simulate a "synthetic Hong Kong" as the counterfactual that is not affected by mass gatherings and public demonstrations. Controlling for other confounding factors that could affect the COVID-19 transmission, such as city-level socioeconomic characteristics and climatic conditions, the synthetic control obtained from this simulation shares similar pre-intervention characteristics with Hong Kong. By comparing the actual number of confirmed cases with the estimated outcomes in our counterfactual model, we can empirically assess the impact of mass gatherings on the transmission of COVID-19 in Hong Kong.

Mass gatherings, especially public gatherings and demonstrations re-emerged after various social distancing policies, including the prohibition on group gathering regulation (POGGR), were lifted in June. These mass gatherings have a higher level of social mixing and thereby pose a higher risk of virus spreading, compared to routine activities (e.g., school, work, shopping, entertainment, etc.) and social gatherings. Our *hypothesis* is that mass gatherings happened after the relaxation of social policy contributed to the third wave of COVID-19 outbreaks in Hong Kong since early July.

Small-size gatherings (less than 50 people) after the relaxation of POGGR on June 19 may also confound the influence of mass gatherings on the transmission of COVID-19. In our empirical model, we apply two approaches to rule out this confounding effect. First, the potential donor units (i.e., all mainland cities) in our SCM model also adopted restrictions on public gatherings during the entire study period. According to the guidelines announced by the Chinese authorities, public gatherings were to be avoided in principle for medium and

high-risk areas and only be held prudently in regions with low-level risk⁸. All urban and rural communities were to strictly manage community activities and restrict large-scale public gatherings until the end of the pandemic⁹. In fact, as of July 2020, mass gatherings of more than 200 people were still prohibited in most mainland cities. Therefore, the synthetic control constructed from these donor units can rule out the confounding effects of the relaxation of POGGR (up to 50 people as of June 19) because it theoretically allows for mass gatherings of more than 50 people (normally up to 200 people), yet Hong Kong only allows for gatherings of 50 people. If Hong Kong still exhibits a higher (and increasing) infection rate than the synthetic control during the intervention period, it is unlikely to have been caused by these small-size public gatherings. Second, we conducted robustness tests by changing the intervention point in the simulation model. Specifically, we use each of the 10 days following June 20 as the intervention point (i.e., June 20, 21, 22...), which gives us a longer control period to partially account for the impact of the relaxation of POGGR on June 19. Allowing for a longer control period that also takes into account the incubation period and the testing lags in COVID-19 detection.

Another advantage of using a longer control period is that the model can evaluate the effectiveness of those social distancing policies implemented until mid-June. In fact, the daily reported number of new cases of COVID-19 in Hong Kong had stabilized and remained low (almost 0 local cases in a single day, see Fig 1) until the beginning of July. This suggests the various social distancing policies implemented before the occurrence of the large-scale public gatherings and demonstrations in June and July were effective in containing local transmission. The empirical model elaborated below can more accurately identify any nodes of exceptional increases in daily new confirmed cases during the control period.

Nonetheless, we recognize that this study has significant limitations. There are disparities in the social and political environment of Hong Kong and mainland cities, for example, Hong Kong has a unique political system and economic autonomy, and different levels of citizen compliance to government policies as well as public trust in government. Such embedded discrepancies might affect our estimation of the impact of mass gatherings' impact on COVID-19 transmission. While acknowledging the importance of capturing these factors in reproducing the synthetic control, we are not able to control them accurately in the SCM model due to their time-invariant nature. Instead, we introduce an extended pre-intervention period in our model and simulate the synthetic control with a border span of predictors over that period. Therefore, the “synthetic HK”, which is constructed by the optimal weighted combination of potential donor cities, is simulated with the greatest possible matching for all predictors and variables of interest, and any unobservable factors can be implicitly captured variables.

⁸ Available at

http://english.www.gov.cn/policies/latestreleases/202004/08/content_WS5e8dd995c6d0c201c2cc07a7.htm

⁹ http://www.gov.cn/xinwen/2020-01/30/content_5473104.htm

Data and data sources

This research aggregates a COVID-19 database with city-level socioeconomic, demographic, and climatic data for more than 200 Chinese cities to construct a counterfactual context for comparative analysis to examine the impact of Hong Kong's mass gatherings, public demonstrations in particular, on the local outbreak of COVID-19. The existing literature suggests that the development of pandemics is catalyzed by multiple factors. Therefore, we incorporate a number of predictors into our SCM models, including epidemiological variables (i.e., the number of new/total confirmed cases), city-level basic characteristics, and climatic parameters. Additionally, three epidemiological variables are also included as predictors to estimate the trends of infection in the 14-day before intervention and preclude any confounding factors associated with the systematic discrepancies between the simulated synthetic control and HK. They are, the average of total case numbers, average of total cases per 10,000 persons (taking population size into consideration), and sum of daily new cases, which are calculated for the 14-day period before the intervention point. Our SCM models take the indicators of new and total cases daily to capture the outbreak situation scenario and use other city-level variables as predictors. All the city-level data is gathered from several sources.

City-level COVID-19 data

The primary data of our SCM analysis is the city-level daily epidemiological data of the COVID-19 outbreak. A China Data Lab collected the daily epidemiological data between June and July and the data are openly accessible on Harvard Dataverse. We retrieved data from this dataset on the daily number of new and total cases for each prefecture-level city in the mainland, as well as for Hong Kong and Macau.

Demographic and socioeconomic data

A number of studies have suggested that regional socioeconomic factors, such as the concentration level of population, total factor productivity, the degree of unemployment, accessibility to healthcare influence how fast and how far infectious diseases are transmitted (Tuckel et al., 2006; Mamelund et al. 2013; Grantz et al., 2016; Adda, 2016; Clay et al., 2019; Li et al., 2021; Zhu & Guo, 2021; Zhou et al., 2022; Zhu et al., 2022). It is, therefore, necessary to consider demographic characteristics (average population, number of households, population density). Demographic statistics of each city were derived directly from the latest version of the China City Statistical Yearbook (2019), which provides information in 2018. To calculate the population density of these cities, we utilized geographic data on the total land area from the same data source. For data on Hong Kong and Macau, we used 2018 demographic data from the official Statistical Yearbooks provided by their Socioeconomic and Statistics Departments. We incorporated several predictor variables to control for socioeconomic factors in our empirical model, including the level of economic development (represented by total GDP, per-capita GDP) and public health capacity (represented by the number of hospitals, and the numbers of doctors and hospital cots per 10,000 people).

Natural meteorological parameters

Previous research has demonstrated that weather conditions and air quality potentially play a role in the spread of COVID-19 (Bannister-Tyrrell et al., 2020; Chen et al., 2020; Jia, Ding et al., 2020; Shi et al., 2020; Wang, Jiang et al., 2020; Tosepu et al., 2020; Zhu et al., 2022) and other infectious diseases (Ye, Zhang et al., 2004; Tan et al., 2005; Bi, Wang & Hiller, 2007; Clay et al., 2015; Clay et al., 2019). Hence, this research considers various daily natural meteorological factors, including air quality index (AQI), average temperature, wind speed, and relative humidity, to simulate the counterfactual scenario and estimate case numbers.

These factors are all proven to impact the transmission of an epidemic in the previous literature, but many studies on COVID-19 transmission do not control for their effects. Simulated by the SCM models, the constructed synthetic control in the absence of mass gatherings has similar climate conditions to Hong Kong; therefore, we are able to control for any influence on the estimated case number of COVID-19.

Climatic parameters for cities in mainland China were derived from the China Meteorological Data Service Centre, which provides hourly records of common meteorological factors such as temperature, relative humidity as well as wind speed at each meteorological observatory station. We calculated daily numbers by averaging the hourly data for each variable. For each city with epidemiological data, we used an Empirical Bayesian Kriging interpolation to compute the values of the three meteorological elements in ArcGIS. Kriging is a statistical approach for predicting the best possible location in a geographic region, which has been widely used in meteorological applications, agriculture, geoscience, and many other disciplines because of its minimized prediction error. Because it offsets the deviations resulting from the semivariogram model, the Empirical Bayesian Kriging method is more resilient than conventional Kriging techniques (Krivoruchko, 2012). Therefore, we adopt the Empirical Bayesian Kriging as the interpolation method of meteorological parameters in this study; the parameters in the model are the default settings of the software. The daily climatic parameters of Hong Kong are collected from the Hong Kong Observatory, and those of Macau SAR are extracted from the Meteorological and Geophysical Bureau.

We captured the dynamics of air quality in each city by Air Quality Index (AQI), which is measured based on the number of six atmospheric pollutants (e.g., SO₂, PM_{2.5}, PM₁₀) detected at all inspection stations within the city border. The source of Chinese cities' air quality data is from Harvard Dataverse. Each record includes information on the daily average, maximum, minimum, and standard deviation value of AQI. In this study, we use the daily average AQI to represent the air quality conditions of each city. Meanwhile, World Air Quality Index (WAQI) project provides daily AQI data of all air quality monitoring stations in Hong Kong and Macau. The values of daily AQI in Hong Kong (or Macau) were calculated by averaging the daily data of all air quality monitoring stations in the region.

Model Specification

Through optimal weighted linear simulations using data from over 200 Chinese cities, we simulate a "synthetic Hong Kong" that has many of the same features as the actual Hong Kong. We estimate COVID-19 transmission patterns using such synthetic control for a counterfactual scenario in which there were no public gatherings and demonstrations in Hong Kong. Based on aggregated epidemiological data from chosen cities, we estimate the trend in the daily numbers of new and total cases in this counterfactual circumstance (i.e., donor units). We may analyze the influence of mass gatherings on COVID-19 transmission in local communities in Hong Kong by comparing the real epidemiological trajectory recorded in the city with simulated trends, and therefore measure the effectiveness of the HK government's social distancing regulations.

As discussed above, our SCM model uses a long control period starting on May 1 and ending on June 19 to eliminate the confounding effects of all other events that happened during the same period as the public gatherings and demonstrations studied. As the first large-scale

political gathering in our study period happened in Hong Kong on June 20, the beginning of the intervention period is designated as June 20. During the intervention period, a series of public gatherings and demonstrations occurred (see Fig 2). We expect that significant differences in infection numbers between the synthetic control and Hong Kong may not be observed for days immediately following June 20. This is because of the existence of a median incubation period of 5 days and a potential testing and reporting lag of 1-2 days. People infected during public gatherings and demonstrations may not be reported as confirmed cases until several days after the protest. Taking into account these potential lags, we also conduct robustness checks by using each of the 10 days following June 20 as different intervention points (i.e., June 21, 22, 23...).

The SCM method has been widely used to assess the impact of political reforms and various economic policies. (Abadie & Gardeazabal 2003; Billmeier & Nannicini, 2013; Abadie et al. 2010; Pieters et al., 2016; Barlow et al., 2017; Adhikari et al., 2018; Marrazzo & Terzi, 2017). Rather than selecting a real city as the counterfactual for comparison, our SCM combines several comparison units (i.e., over 200 Chinese cities) from the donor city pool by an optimum weight of each donor cities, to construct one a synthetic Hong Kong as counterfactual. The inherent advantage of the SCM is that it offers a data-driven method for determining a matching counterfactual used for comparative analysis. As a result, issues of ambiguity, endogeneity, and self-selection biases can be eliminated.

The following is an illustration of our SCM model: Y_{jt} is defined as COVID-19 infections diagnosed in city j at time t . There are $J+1$ cities, with Hong Kong as the first city ($j = 1$) being subject to intervention (i.e., the occurrence of mass gatherings in the form of public gatherings and demonstrations) after time T_0 (i.e., June 20), and the remaining J cities (selected cities in China) are donor cities for simulating the “synthetic Hong Kong”. The number of cases for city j with the occurrence of mass gatherings is set as Y_{jt}^I and the estimated infection trend in the synthetic control is set as Y_{jt}^N , the later hypothetically assume that mass gatherings had not occurred at time $t \in [T_0 + 1, T]$. Thus, Y_{jt}^I is observable as

$$Y_{jt}^I = Y_{jt}^N + \tau_{jt}D_{jt}$$

where τ_{jt} is the impact of mass gatherings on COVID-19 cases number for city j at time t (i.e., from June 20 onwards). As Hong Kong, that is, $j = 1$, is the only city that affected by mass gatherings, D_{jt} is the indicator written as the following:

$$D_{jt} = \begin{cases} 1 & \text{for } j = 1, t > T_0 \\ 0 & \text{otherwise} \end{cases}$$

D_{jt} indicates whether city j was influenced by public gatherings in period t (i.e., from June 20 onwards). D takes value 1 for the treated unit ($j=1$), and 0 otherwise. For $t > T_0$, $\tau_{1t} = Y_{1t}^I - Y_{1t}^N$ gives the net effect of mass gatherings on the increase of COVID-19 infections.

The objective is to assess the value of $\tau_{1T_0+1}, \tau_{1T_0+2}, \dots, \tau_{1T}$.

In accordance with the algorithm stated in Abadie et al. (2010), we assume that Y_{jt}^N is given by a factor model

$$Y_{jt}^N = \delta_t + \theta_t Z_j + \lambda_t \mu_j + \varepsilon_{jt}$$

where $Z_j = [Z_{j1}, \dots, Z_{jr}]'$ is a $(r \times 1)$ vector of predictors that are observable and irrelevant to the occurrence of mass gatherings, as presented in Table 2. $\theta_t = [\theta_{1t}, \dots, \theta_{rt}]$ is a $(1 \times r)$ vector of parameters. Also, a $(1 \times k)$ vector of common factors which are unobservable is captured by $\lambda_t = [\lambda_{1t}, \dots, \lambda_{kt}]$. A $k \times 1$ vector of unknown factor loadings is represented by $\mu_j = [\mu_{j1}, \dots, \mu_{jk}]'$. Unobservable ephemeral impacts with zero mean are depicted by the error terms ε_{jt} . A constant factor loading is described by δ_t , a common factor which is unknown.

The actual case number reported in Hong Kong are observable and we have $Y_{jt}^l = Y_{1t}$, but we have no way to observe how many infected cases would be under the conditions of the synthetic control Y_{1t}^N within the period from $T_0 + 1$ to T . In order to approximate the value of Y_{1t}^N , we utilized the procedure stated in Abadie & Gardeazabal (2003). A $(J \times 1)$ vector of weights $W = [w_2, \dots, w_{J+1}]'$ is simulated so that we can have $\sum_{j=2}^{J+1} w_j = 1$, and $w_j \geq 0$ for $j = 2, \dots, J + 1$.

By using a weighted average of donor cities with observed outcomes and covariates among the time period before the intervention point, we estimate Y_{1t} and Z_1 for the phase before the occurrence of intervention, i.e., $t \leq T_0$. Thus, $W = (w_2, \dots, w_{J+1})$ is computed for $t \in [1, T_0]$, such that

$$Y_{1t} = \sum_{j=2}^{J+1} w_j Y_{jt}$$

$$Z_1 = \sum_{j=2}^{J+1} w_j Z_j,$$

Once W is estimated, the effect of intervention (i.e., mass gatherings) at $t = T_0 + 1, T_0 + 2, \dots$ can be captured by

$$\tau_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j Y_{jt}$$

W is the optimum weight for the combination of control units (i.e., donor cities), which minimizes the disparities of predictor and outcome variables between the reality and the counterfactual among the pre-intervention phase. We utilized the technique described by Abadie et al. (2010) to get the optimal weights, which technically minimizes the mean squared prediction error of all predictors and outcome variables over the pre-intervention period.

Results

After data cleaning, 282 Chinese cities were included in the donor pool for constructing the synthetic control. In our baseline model in Panel A, we designate June 20, one day after the lift of POGGR, as the beginning of the intervention period, when Hong Kong began observing mass gatherings including a series of public gatherings and demonstrations. Table 1 demonstrates the distribution of sample weights of the selected comparison cities for the synthetic control simulated in Panel A. We constructed the synthetic control and estimated the trends of the daily numbers of COVID-19 new and total cases in Model 1 and Model 2. The synthetic control obtained in Model 1 was constructed using 6 donor cities, while that obtained in Model 2 was from 4 donor cities. Table 2 demonstrates the balance of predictor variables during the pre-intervention period, indicating that the characteristics of the synthetic control are very similar to those in Hong Kong before the occurrence of public gatherings and demonstrations. This suggests that our synthetic control simulated from the SCM models is a good fit with the real Hong Kong in the pre-intervention period.

Table 1: Distribution of sample weights in the donor pool for synthetic Hong Kong

| Panel A: intervention point on June 20 | | | |
|---|--|--|--|
| Model 1 (Fig 3) | | Model 2 (Fig 4) | |
| Variable of interest: new cases | | Variable of interest: total cases | |
| Name of the donor city | Weight in the synthetic control | Name of the donor city | Weight in the synthetic control |
| Jilin | 0.265 | Jilin | 0.47 |
| Suihua | 0.261 | Shanghai | 0.427 |
| Shanghai | 0.221 | Beijing | 0.192 |
| Macau | 0.159 | Wuhan | 0.089 |
| Beijing | 0.082 | | |
| Huanggang | 0.012 | | |

Table 2: Pre-intervention balance of predictor variables in Panel A (June 20 as the intervention point)

| Pre-intervention characteristics during the control period (May 1 – June 19) | | | |
|---|----------------|---------------------------------|---------------------------------|
| | Real HK | Model 1 Synthetic HK | Model 2 Synthetic HK |
| # of new cases in the past 14 days | 22.14 | 6.90 | 12.5 |
| Preintervention 14-day (June 6- June 19) average of daily total case numbers | 1,111 | 271.81 | 1,078.14 |
| Preintervention 14-day (June 6- June 19) average of daily number of total cases per 10,000 people | 1.49 | 1.47 | 1.18 |
| Preintervention 14-day (Jun 6- Jun 19) sum of new confirmed cases | 25 | 19.43 | 24.83 |
| Region GDP (million yuan) | 2,398,046 | 1,069,235 | 1,750,217 |

| | | | |
|---|----------|----------|---------|
| # of persons per sq.km (population density) | 6,736.89 | 4,238.87 | 1295.06 |
| # of Hospital Beds per 10,000 persons | 54.27 | 59.83 | 90.98 |
| # of Doctors per 10,000 persons | 19.66 | 39.12 | 52.75 |
| Average Temperature (°C) | 28.31 | 20.18 | 20.03 |
| Relative Humidity (%) | 81.92 | 71.30 | 72.29 |
| Wind Speed (m/s) | 5.29 | 2.76 | 2.68 |
| Air Quality Index (AQI) | 44.28 | 45.21 | 52.03 |

Fig 3 shows the comparison between the actual and simulated trends of COVID-19 new cases confirmed on a daily basis. The residual between the two trajectories after the date of intervention can be interpreted as the impact of mass gatherings after the relaxation of social distancing order, i.e., POGGR, on COVID-19 transmission in local communities. As shown in the figure, Hong Kong saw a much higher number of daily new infections than the synthetic control that models a hypothetical situation assuming no mass gathering happened. This indicates that public gatherings and demonstrations in Hong Kong contributed to the increase in new cases on a daily basis. In the absence of public gatherings and demonstrations, Hong Kong would have no more than 2 new cases registered on any single day from June 20 until the end of July. To explain this in another way, we can calculate the sum of new infections by adding up these predicted daily new cases. For the first 14 days after the intervention date, Hong Kong would have observed around 16 new cases if the POGGR had not been lifted, while the actual number of new cases was 120 during this period. After the first two weeks, the outbreak was under control in the synthetic control, with at most 1 new case confirmed in any single day throughout the rest of July. However, the actual infection numbers in Hong Kong continued to rise dramatically, and as many as 2145 new cases were recorded from June 20 to July 31, significantly more than the 23 new cases estimated in the synthetic control during the same period. This confirms that the protest and demonstration activities in late June and early July contributed to a considerable growing trend of daily new cases.

Fig 3: Impact of mass gatherings (after the lift of POGGR) on COVID-19 new cases

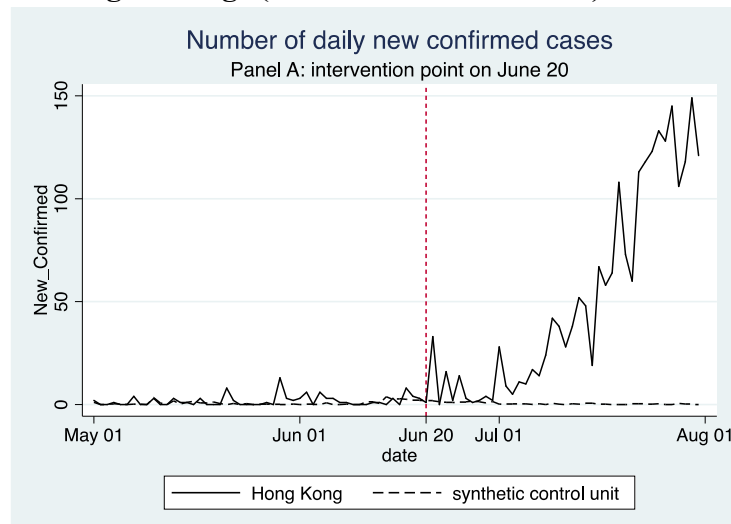
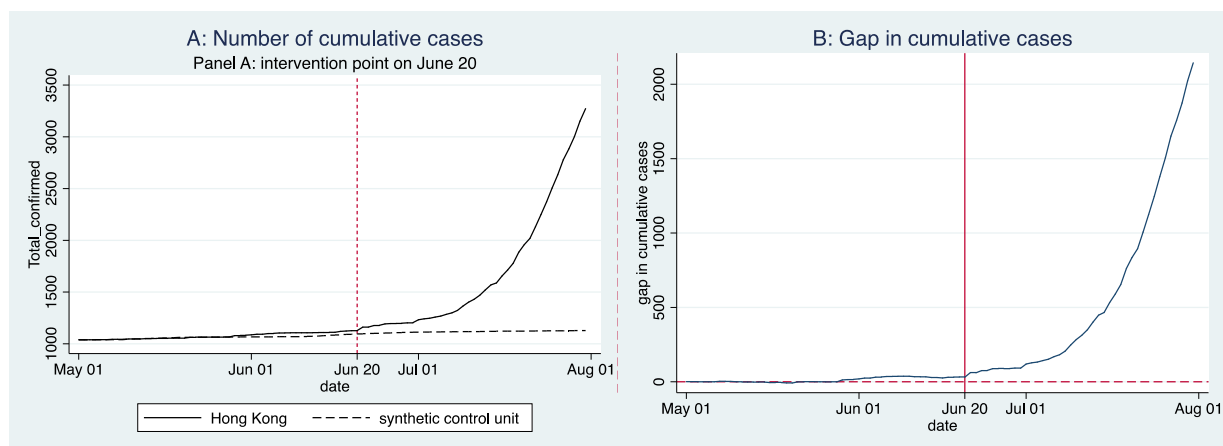


Fig 4A shows the trends of total infections reported and estimated in our synthetic control. The visual comparison suggests that a significant increase in the number of cumulative infections may be attributed to public gatherings and demonstrations during the period from June 20 to the end of July. Before the intervention point (June 20), the figure of total infections in both scenarios (i.e., with and without mass gatherings) is similar, but after June 20, the gap between the two lines widens sharply. This gap is also represented as a line graph in Fig 4B, indicating that public gatherings generally increase COVID-19 infections in the local communities of Hong Kong. The results suggest that if Hong Kong had not been affected by public gatherings and demonstrations after June 20, the outbreak would have stayed at a moderate level, with just around 1128 total cases by July 31, while actual infections exceeded 3000 cases. More importantly, the actual growth rate of infections in the city is significantly higher than that of the counterfactual model. During the period from June 20 to July 31, the total number of infections increased by 2144 cases in Hong Kong, and the average daily growth rate of total cases¹⁰ was 2.63%. In our counterfactual model, however, there was an increase of only 32 cases in the number of total infections; the average daily growth rate was a mere 0.07%. Our results suggest that, if Hong Kong had not been affected by mass gatherings after June 20, the city could have controlled the increase of new infections and kept the number of total infections at a low level.

Figure 4: Impact of mass gatherings (after the lift of POGGR) on COVID-19 total cases

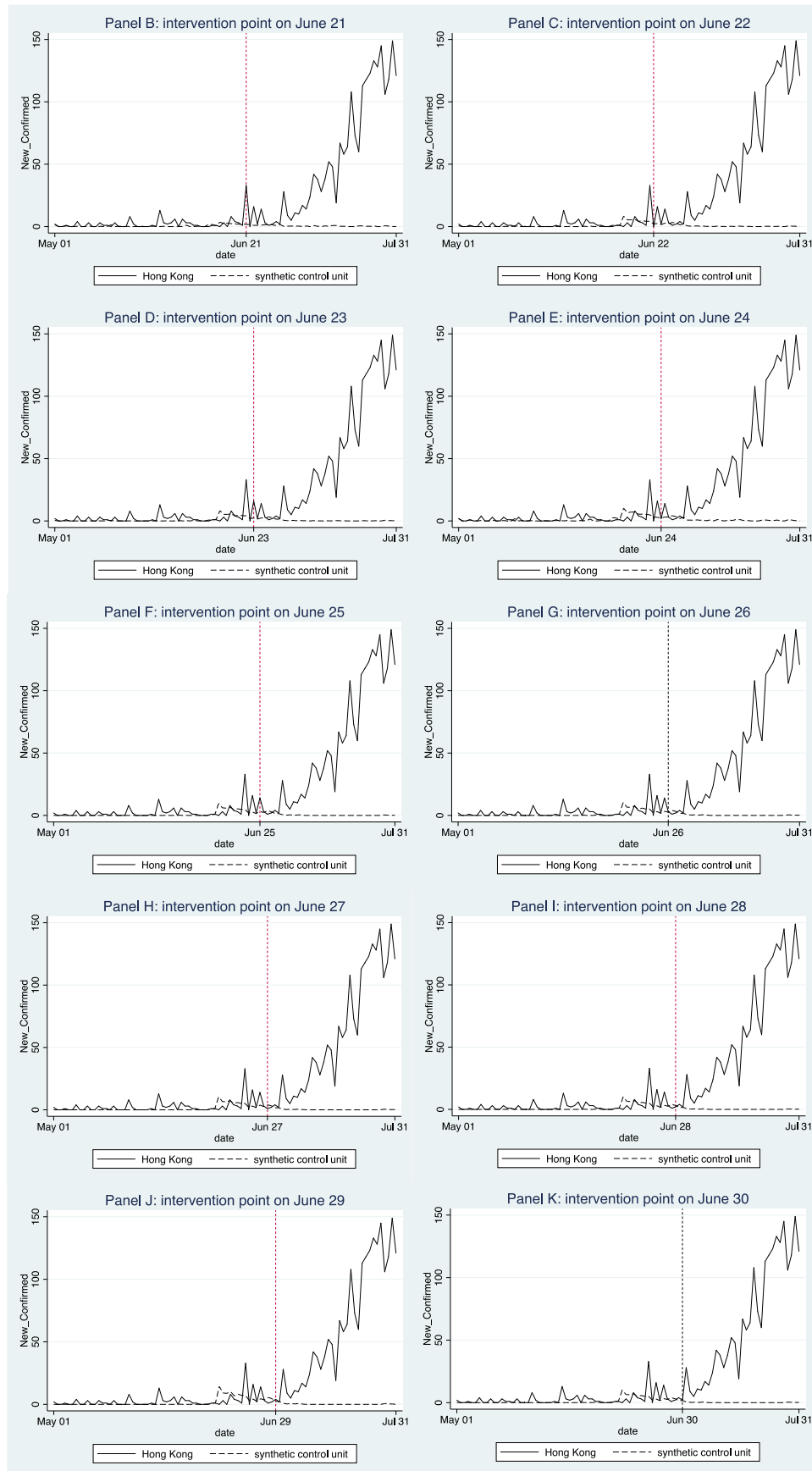


To cover a longer control period and to consider the long incubation period of COVID-19 (i.e., delayed symptom onset after infection), as well as delays in testing and reporting, we changed the intervention point to conduct robustness tests. Here we focus on Model 1, which simulates the trend in daily new cases. If the results in Model 1 are robust, we can infer that the results of Model 2 are also robust. Using the same predictors and outcome variable, we

¹⁰ Calculating the average daily growth rate of total infections: Hong Kong has 1128 total cases on June 20 compared to an estimated number of 1125 total cases in the synthetic control. On July 31, Hong Kong counts 3272 cases while the synthetic control has 1164 cases, in total. From June 20 to July 31, The average daily growth rate in Hong Kong is denoted by r and can be computed from $1128(1+r)^{41}=3272$. Similarly, the average daily growth rate in the synthetic unit rc can be computed from $1125(1+rc)^{41}=1164$.

use each of the 10 days after June 20 as the intervention point (i.e., June 21, 22, ..., 30) in each panel to re-examine our estimation model. Fig 5 shows the results obtained in each Panel with different intervention points. The sample weights of comparison cities that construct the synthetic Hong Kong are provided in the Appendix.

Fig 5: Impact of mass gatherings (after the lift of POGGR) on the number of new cases of COVID-19 (with intervention point changed)



When we changed the intervention point (i.e., the starting date of lifting of public gathering restrictions in Hong Kong), we obtained different estimates of new cases in the respective synthetic controls. However, the adjusted results still show that the actual number of daily new cases of COVID observed after the lift of public gatherings and the occurrence of and demonstrations is, in general, higher than the counterfactual estimated in the synthetic controls obtained in all the Panels. Except for Panels I, J, and K, other Panels using an earlier intervention point indicate that the counterfactual estimates of new case numbers are initially close to the actual number. After some fluctuations, the synthetic trends stabilize in the later stage, while the actual infections numbers grew sharply. If we impose longer lags (as in Panel I, J, and K), meaning that we have a later intervention point, the gap between infection numbers under the two scenarios becomes more significant even in the early stage of the intervention.

Table 3 summarizes the magnitude of the impacts during the 10-day and 30-day periods after the intervention point. In our baseline configuration of the SCM model (Panel A), which uses the start of protest activities (June 20) as the intervention point, the size of the impact of mass gatherings in the first 10-day period after the POGGR relaxation is a growth of 62 new cases. The 30-day impact is an increase of 737 new cases. These increases in new cases due to the impact of mass gatherings account for 87.58% and 97.23% of the total new cases accumulated over the respective time periods. We also conduct robustness checks for these results by extending the control period to capture the lags in the onset of symptoms and testing delays, where we use each of the 10 days after June 20 as the intervention point. When considering shorter lags (1-5 days), we estimate that, on average, protest activities contribute to an increase of around 57 new infections (or 75% of total new cases) over the 10-day post-intervention period and an increase of around 961 new infections (or 97.66%) over the 30-day period. When we consider longer lags (6-10 days), the impact is more profound, with an increase of 101 new cases (or 91.82%) and 1553 new cases (or 99.36%) over the 10-day and 30-day periods, respectively. Therefore, we conclude that public gatherings and demonstrations in Hong Kong in late June and early July contributed significantly to the new wave of COVID-19 infections that began in early July.

Table 3: Summary of the impact of mass gatherings (after the relaxation of social distancing policy) on the increase of new cases during the 10-day and 30-day periods after intervention, taking June 20 and each of the 10 days afterward as the intervention point

| Baseline | | | Robustness tests | | | | | | | | | |
|----------------------------------|-------------------|-----|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Panel A | | | Panel B | Panel C | Panel D | Panel E | Panel F | Panel G | Panel H | Panel I | Panel J | Panel K |
| Intervention point | | | June 21 | June 22 | June 23 | June 24 | June 25 | June 26 | June 27 | June 28 | June 29 | June 30 |
| 10-day period after intervention | Hong Kong | 76 | 77 | 72 | 81 | 70 | 79 | 75 | 89 | 102 | 124 | 162 |
| | Synthetic control | 14 | 13 | 21 | 18 | 24 | 17 | 15 | 13 | 9 | 8 | 4 |
| | Difference | 62 | 64 | 51 | 63 | 46 | 62 | 60 | 76 | 93 | 116 | 158 |
| 30-day period after intervention | Hong Kong | 758 | 830 | 857 | 970 | 1072 | 1193 | 1312 | 1437 | 1581 | 1685 | 1799 |
| | Synthetic control | 21 | 20 | 25 | 22 | 34 | 17 | 16 | 13 | 9 | 8 | 5 |
| | Difference | 737 | 810 | 832 | 948 | 1038 | 1176 | 1296 | 1424 | 1572 | 1677 | 1794 |

Discussion

This study investigates whether the relaxation of social distancing policies and the subsequent occurrence of public gatherings contributed to a new wave of local outbreak (also known as the third wave) in Hong Kong. After more than two years into the pandemic, the global scientific community has gained significant understanding of COVID-19 regarding the relationship between small group gatherings and viral transmission, especially in indoor settings. However, previous studies leave a research gap in terms of how mass gatherings in public places, which generally have lower perceived risks but higher uncertainties, could influence local transmission. By examining the series of public demonstrations that happened in Hong Kong following its relaxation of the Prohibition On Group Gathering Regulation (i.e., the POGGR), this study provides important empirical evidence for the implementation of social distancing policies in limiting mass gatherings and controlling virus transmission.

Various public gatherings and demonstrations took place in Hong Kong from June 20, 2020 onwards. These events were collectively influenced by the enactment of the NSL in the city and the coincidental relaxation of POGGR. Although people usually were wearing masks during public gatherings and demonstrations, these activities still tend to be highly crowded with increased physical contact, creating openings for virus transmission and increasing the rate of COVID-19 infection. Some have suspected that the third wave was a result of the loosened quarantine requirements, especially the quarantine exemptions for special overseas arrivals such as seafarers.^{11,12} However, the majority of cases during the investigation period (over 83%) were reported as local infections¹³, implying that the lifting of social distancing order played a role in influencing the transmission during Hong Kong's third wave. Using an aggregated dataset of epidemiological, city-level meteorological and socioeconomic data, our SCM modeling results clearly evidence that public and social mass gatherings that occurred after the relaxation of social distancing policy had a sizable impact on COVID-19 transmission in Hong Kong's local communities. Our findings indicate that from June 20 to July 31, 2020, Hong Kong observed an infection rate, as captured by the mean daily growth rate of total cases, substantially greater than that estimated by the counterfactual control unit (2.63% versus 0.07%). This transmission rate has led to exponential growth in new cases. We also estimate that public gatherings and demonstrations increase the number of new infections by 62 cases (or 87.58% of total new cases) over the 10-day period after they first occurred on June 20 and by 737 cases (or 97.23 %) over the 30-day period. These findings are in line with local clinical experts who argued that the propagated outbreak in the third wave derived from increased social gatherings during dragon boat festival (June 25) and July 1 holiday (Sridhar, 2020).

¹¹ Government anti-epidemic advisers and medical experts estimate that the new wave of outbreaks originated from a number of people who were exempted from quarantine and introduced into the community via taxi drivers.

¹² What Hong Kong can do to avoid a total lockdown and a fourth wave of Covid-19 cases. South China Morning Post.

https://sc.mp/k852a?utm_source=copy_link&utm_medium=share_widget&utm_campaign=3095460

¹³ Details of previous COVID-19 cases in Hong Kong, reported by the Centre for Health Protection (CHP) of the Department of Health (DH). https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf

Based on our aforementioned findings, this research provides important guidance for *local anti-pandemic policymaking in Hong Kong*. We recommend that it is vital for the Hong Kong government to strengthen its law enforcement to caution any large-scale public gatherings during the pandemic, including public gatherings and demonstrations. Our findings suggest that participants in large-scale gatherings are susceptible to infection when the virus has not been completely controlled. Based on Hong Kong's experience during the third wave, the relaxation of social distancing restrictions in late June 2020 created a loophole for large-scale public gatherings. In particular, the lifting of the POGGR on June 19 created opportunities for small- or medium-sized gatherings and even large-sized demonstrations (though unpermitted) to occur. Our empirical findings corroborate that these gatherings have led to a resurgence in local COVID-19 infections. Also, *social distancing policies are of great importance in limiting mass gatherings and managing the risk of local transmission of COVID-19 when herd immunity has not yet gained*.

Hence, in this third wave of Hong Kong's local outbreak, it was crucial to tighten restrictions on mass/group gathering promptly to curb the transmission of potential community infections and alleviate pressure on the public health system in Hong Kong. In fact, in August 2020, Hong Kong tightened the POGGR to a limit of 2 people, and later lifted this to 4 people in September 2020. As a result, the third wave gradually came under control. Acknowledging the uncertainty of the COVID-19 pandemic and the unique political and social challenges in Hong Kong, the government departments should commit to containing the pandemic for public society and welfare. The health department should attentively assess and monitor risks stemming from any mass gatherings and prompt adjustment of social distancing policies, accordingly. Public leaders of other functional constituencies should cooperate to address the root causes of social unrest in society, improve social welfare and rebuild trust in government among citizens. In such a way, the effectiveness of the government's anti-pandemic policies can be guaranteed.

Our findings on the significant impact of mass gatherings on the transmission of COVID-19 in Hong Kong also provide strong evidence for the necessity of implementing social distancing policies *in other parts of the world, especially in places where large-scale gathering events are also being experienced*. The tragedy of the deadly 2022 Halloween Crush in South Korea has made the entire nation mourn, yet its immediate impact on the possible eruptive COVID-19 new cases may cause the nation to suffer much graver losses. In the face of other novel infectious diseases (NID) in the future, social distancing policies, and particularly those restrictions on group gatherings, should be prioritized once the local outbreaks are observed, and should be maintained until the outbreak is completely curbed. Constantly changing social distancing orders will create confusion and inconvenience among the public, as well as hinder the effectiveness of prior social distancing efforts.

In conclusion, this research not only makes significant contributions to the scientific literature on the impact of mass gathering on the transmission of COVID-19, but also has important implications for policymakers in designing more effective policies to control the pandemic.

First, by paying attention to public gatherings that often take place in outdoor environments where ventilation is usually better, our research can help complement existing empirical studies that focus on how gatherings in indoor settings may influence virus transmission. The findings of this study will help advance our understanding of the transmission patterns and sources, and thereby provide scientific guidance for managing the risks of mass gathering events. *Second*, the study in Hong Kong not only helps the local government improve its policymaking capacity to combat the pandemic, but also contributes to relieving pandemic situations in other places around the world that are also experiencing mass gatherings (e.g., large-scale religious events). Hence, our findings can generate useful insights on pandemic control for countries considering or starting to relax their restrictions on group gatherings to adapt to the new normal during a pandemic.

Lastly, it is noted that this study does not address the more transmissible SARS-CoV-2 variants that were later observed in multiple countries. This is because the earliest SARS-CoV-2 variant, i.e., Alpha, first appeared in the U.K. in November 2020¹⁴ and our study period is restricted to between May and July 2020. Moreover, our study period also does not reflect on the ongoing situation wherein many cities/countries in the world have gained substantial immunity to COVID-19. Nevertheless, the past anti-pandemic experiences are still important lessons for the future. This study examines how public gatherings can influence the transmission of contagious viruses when herd immunity is yet to form. This can help to provide crucial evidence for evaluating the use of social distancing policies against early-stage outbreaks of other novel infectious diseases in the future. Such evidence is pivotal for both the scientific community and the general public.

¹⁴ For more information about the evolution of the SARS-CoV-2 virus, see WHO's official declaration at <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> and Yale Medicine's guide to the coronavirus variants at <https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron>.

Reference

- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*. <https://doi.org/10.1257/000282803321455188>
- Abadie, A., Diamond, A., & Hainmueller, A. J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco control program. *Journal of the American Statistical Association*.
<https://doi.org/10.1198/jasa.2009.ap08746>
- Abubakar, I., Gautret, P., Brunette, G. W., Blumberg, L., Johnson, D., Pomeroy, G., ... & Khan, A. S. (2012). Global perspectives for prevention of infectious diseases associated with mass gatherings. *The Lancet infectious diseases*, 12(1), 66-74.
- Adda, Jerome (2016). Economic activity and the spread of viral diseases: evidence from High frequency data. *Q. J. Econ.* 131 (2), 891–941.
- Adhikari, B., Duval, R., Hu, B., & Loungani, P. (2018). Can Reform Waves Turn the Tide? Some Case Studies using the Synthetic Control Method. *Open Economies Review*.
<https://doi.org/10.1007/s11079-018-9490-3>
- Ai, S., Zhu, G., Tian, F., Li, H., Gao, Y., Wu, Y., ... & Lin, H. (2020). Population movement, city closure and spatial transmission of the 2019-nCoV infection in China. *medRxiv preprint*. Available on doi: <https://doi.org/10.1101/2020.02.04.20020339>.
- Al-Tawfiq, J. A., & Memish, Z. A. (2012). Mass gatherings and infectious diseases: prevention, detection, and control. *Infectious Disease Clinics*, 26(3), 725-737.
- Al-Tawfiq, J. A., Memish, Z. A., & Zumla, A. (2021). Mass Religious Gatherings Events and COVID-19-Easing of COVID-19 Restrictions and a Staged Approach to Scaling up the Umrah Pilgrimage.
- Bajardi, P., Poletto, C., Ramasco, J. J., Tizzoni, M., Colizza, V., & Vespignani, A. (2011). Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLoS ONE*, 6(1), e16591.
- Bannister-Tyrrell, M., Meyer, A., Faverjon, C., & Cameron, A. (2020). Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to February 29 2020. *medRxiv preprint*. Available from:
<https://doi.org/10.1101/2020.03.18.20036731>
- Barlow, P., McKee, M., Basu, S., & Stuckler, D. (2017). Impact of the North American free trade agreement on high-fructose corn syrup supply in Canada: A natural experiment using synthetic control METHODS. *CMAJ*. <https://doi.org/10.1503/cmaj.161152>
- Barrios, J., Benmelech E., Hochberg Y., Sapienza P. & Zingales L. (2020). Civic Capital and Social Distancing during the Covid-19 Pandemic, NBER Working Papers 27320, National Bureau of Economic Research, Inc.
- Bi, P., Wang, J., & Hiller, J. E. (2007). Weather: driving force behind the transmission of severe acute respiratory syndrome in china?. *Internal Medicine Journal*, 37(8), 550-554.

- Billmeier, A., & Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*.
https://doi.org/10.1162/REST_a_00324
- Blyth, C. C., Foo, H., van Hal, S. J., Hurt, A. C., Barr, I. G., McPhie, K....Dwyer, D. E. (2010). Influenza Outbreaks during World Youth Day 2008 Mass Gathering. *Emerging Infectious Diseases*, 16(5), 809-815.
<https://doi.org/10.3201/eid1605.091136>.
- Brockmann, D., & Helbing, D. (2013). The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science*, 342(6164), 1337–1342.
- Burke, R. M., Midgley, C. M., Dratch, A., Fenstersheib, M., Haupt, T., Holshue, M., ... & Rolfes, M. A. (2020). Active monitoring of persons exposed to patients with confirmed COVID-19—United States, January–February 2020. *Morbidity and Mortality Weekly Report*, 69(9), 245.
- Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, J. (2016). The Fundamental Determinants of Anti-Authoritarianism. UC-Berkeley, unpublished.
- Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2019). Public gatherings as strategic games: experimental evidence from Hong Kong's anti-authoritarian movement. *The Quarterly Journal of Economics*, 134(2), 1021-1077.
- Charu, V., Zeger, S., Gog, J., Bjørnstad, O. N., Kissler, S., Simonsen, L., ... Viboud, C. (2017). Human mobility and the spatial transmission of influenza in the United States. *PLOS Computational Biology*, 13(2), e1005382.
- Chen, B., Liang, H., Yuan, X., Hu, Y., Xu, M., Zhao, Y., & Zhu, Y. (2020) . Roles of meteorological conditions in COVID-19 transmission on a worldwide scale. medRxiv preprint. Available from: <https://doi.org/10.1101/2020.03.16.20037168>
- Clay, K., Lewis, J., & Severnini, E. (2019). What explains cross-city variation in mortality during the 1918 influenza pandemic? Evidence from 438 U.S. cities. *Economics & Human Biology*, 35, 42–50.
- Clay, K., Lewis, J., & Severnini, E. R. (2015). Pollution, infectious disease, and mortality: evidence from the 1918 spanish influenza pandemic. *Social Science Electronic Publishing*, 114(4), 923–933.
- Cowling, B. J., Lau, L. L., Wu, P., Wong, H. W., Fang, V. J., Riley, S., & Nishiura, H. (2010). Entry screening to delay local transmission of 2009 pandemic influenza A (H1N1). *BMC Infectious Diseases*, 10(1).
- Dave, D., Friedson, A., Matsuzawa, K., Sabia, J. & Safford, S. (2020). "Black Lives Matter Protests, Social Distancing, and COVID-19," IZA Discussion Papers 13388, Institute of Labor Economics (IZA).
- Dave, D., Friedson, A., Matsuzawa, K. & Sabia, J. (2020). When Do Shelter-In-Place Orders Fight COVID-19 Best? Policy Heterogeneity across States and Adoption Time, IZA Discussion Papers 13190, Institute of Labor Economics (IZA).
- Ding, W. , Levine, R. , Lin, C. , & Xie, W. . (2020). Social distancing and social capital: why u.s. counties respond differently to covid-19. NBER Working Papers 27393, National Bureau of Economic Research, Inc.
- Domènech-Montoliu, S., Pac-Sa, M. R., Vidal-Utrillas, P., Latorre-Poveda, M., Del Río-González, A., Ferrando-Rubert, S., ... & Arnedo-Pena, A. (2021). Mass gathering

- events and COVID-19 transmission in Borriana (Spain): A retrospective cohort study. *PloS one*, 16(8), e0256747.
- Ebrahim, S. H., & Memish, Z. A. (2020). COVID-19—the role of mass gatherings. *Travel medicine and infectious disease*, 34, 101617.
- Furuse, Y., Sando, E., Tsuchiya, N., Miyahara, R., Yasuda, I., Ko, Y. K., ... & Oshitani, H. (2020). Clusters of coronavirus disease in communities, Japan, January–April 2020. *Emerging infectious diseases*, 26(9), 2176.
- Gautret, P. (2020). Measles outbreaks at mass gathering mostly occur at youth events. *The Lancet Infectious Diseases*, 20(1), 23.
- Gov. HK (2020). Fighting the Virus for Six Months Battling Another Wave of the Epidemic Together. <https://www.ceo.gov.hk/eng/pdf/article20200725.pdf>
- Grantz, K. H. , Rane, M. S. , Salje, H. , Glass, G. E. , Schachterle, S. E. , & Cummings, D. A. T. . (2016). Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences*, 201612838.
- Gómez-Ríos, D., Ramírez-Malule, D., & Ramírez-Malule, H. (2020). The effect of uncontrolled travelers and social distancing on the spread of novel coronavirus disease (COVID-19) in Colombia. *Travel Medicine and Infectious Disease*, 35, 101699.
- Hoang, V. T., & Gautret, P. (2018). Infectious diseases and mass gatherings. *Current infectious disease reports*, 20(11), 1-12.
- Ishola, D. A., & Phin, N. (2011). Could influenza transmission be reduced by restricting mass gatherings? Towards an evidence-based policy framework. *Journal of epidemiology and global health*, 1(1), 33-60.
- Jia, J., Ding, J., Liu, S., Liao, G., Li, J., Duan, B., ... Zhang, R. (2020). Modeling The Control Of COVID-19: Impact Of Policy Interventions And Meteorological Factors. *Electronic Journal of Differential Equations*, 2020 (23), 1-24.
- Krivoruchko, K. (2012). Empirical Bayesian Kriging: Implemented in ArcGIS Geostatistical analyst. *ArcUser*, 6-10. Available online: <http://www.esri.com/news/arcuser/1012/files/ebk.pdf> (accessed on June 12 2020)
- Ku, C., Ng, T., & Lin, H. (2020). Epidemiological Benchmarks of the COVID-19 Outbreak Control in China after Wuhan's Lockdown: A Modelling Study with An Empirical Approach. *SSRN Electronic Journal*, January 2020.
- Lelieveld, J., Helleis, F., Borrmann, S., Cheng, Y., Drewnick, F., Haug, G., ... & Pöschl, U. (2020). Model calculations of aerosol transmission and infection risk of COVID-19 in indoor environments. *International Journal of Environmental Research and Public Health*, 17(21), 8114.

- Lemey, P., Rambaut, A., Bedford, T., Faria, N., Bielejec, F., Baele, G., ... & Suchard, M. A. (2014). Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2. *PLoS Pathogens*, 10(2), e1003932.
- Li, B., Peng, Y., He, H., Wang, M., & Feng, T. (2021). Built environment and early infection of COVID-19 in urban districts: A case study of Huangzhou. *Sustainable Cities and Society*, 66, 102685.
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... & Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England journal of medicine*.
- Mat, N. F. C., Edinur, H. A., Razab, M. K. A. A., & Safuan, S. (2020). A single mass gathering resulted in massive transmission of COVID-19 infections in Malaysia with further international spread. *Journal of Travel Medicine*.
- Mamelund, S. E., Sattenspiel, L., & Dimka, J. (2013). Influenza-associated mortality during the 1918–1919 influenza pandemic in alaska and labrador. *Social Science History*, 37(02), 177-229.
- Marrazzo, P. M., & Terzi, A. (2017). Wide-reaching Structural Reforms and Growth: A Cross-country Synthetic Control Approach. *Harvard CID Working Paper Series*.
<https://doi.org/10.1111/j.1475-4932.1987.tb00650.x>
- Memish, Z. A., Assiri, A., Turkestani, A., Yezli, S., al Masri, M., Charrel, R., ... & Gautret, P. (2015). Mass gathering and globalization of respiratory pathogens during the 2013 Hajj. *Clinical Microbiology and Infection*, 21(6), 571-e1.
- Murray, C. J., Lopez, A. D., Chin, B., Feehan, D., & Hill, K. H. (2007). Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *Lancet*, 368(9554), 2211-2218.
- Memish, Z. A., Steffen, R., White, P., Dar, O., Azhar, E. I., Sharma, A., & Zumla, A. (2019). Mass gatherings medicine: public health issues arising from mass gathering religious and sporting events. *The Lancet*, 393(10185), 2073-2084.
- Moritz, S., Gottschick, C., Horn, J., Popp, M., Langer, S., Klee, B., ... & Mikolajczyk, R. (2021). The risk of indoor sports and culture events for the transmission of COVID-19. *Nature communications*, 12(1), 1-9.
- Murakami, M., Yasutaka, T., Onishi, M., Naito, W., Shinohara, N., Okuda, T., ... & Imoto, S. (2021). Living with COVID-19: mass gatherings and minimizing risk. *QJM: An International Journal of Medicine*, 114(7), 437-439.
- [Petersen, E., Memish, Z. A., Zumla, A., & Al Maani, A. \(2020\). Transmission of respiratory tract infections at mass gathering events. *Current opinion in pulmonary medicine*, 26\(3\), 197-202.](#)
- Pieters, H., Curzi, D., Olper, A., & Swinnen, J. (2016). Effect of democratic reforms on child mortality: a synthetic control analysis. *The Lancet Global Health*.
[https://doi.org/10.1016/S2214-109X\(16\)30104-8](https://doi.org/10.1016/S2214-109X(16)30104-8)
- Rainey, J. J., Phelps, T., & Shi, J. (2016). Mass gatherings and respiratory disease outbreaks in the United States—should we be worried? results from a systematic literature review and analysis of the national outbreak reporting system. *PloS one*, 11(8), e0160378.
- Russell, C. A., Jones, T. C., Barr, I. G., Cox, N. J., Garten, R. J., Gregory, V., ... Smith, D. J.

- (2008). The Global Circulation of Seasonal Influenza A (H3N2) Viruses. *Science*, 320(5874), 340–346.
- Shi, P., Dong, Y., Yan, H., Zhao, C., Li, X., Liu, W.,...Xi, S.(2020). Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Science of the Total Environment*, 728 (2020) 138890.
- Shi, P., Keskinocak, P., Swann, J. L., & Lee, B. Y. (2010). The impact of mass gatherings and holiday traveling on the course of an influenza pandemic: a computational model. *BMC public health*, 10(1), 1-12.
- Tan, J., Mu, L., Huang, J., Yu, S., Chen, B., & Yin, J. (2005). An initial investigation of the association between the sars outbreak and weather: with the view of the environmental temperature and its variation. *Journal of Epidemiology & Community Health*, 59(3), 186-192.
- Sridhar, S. (2020). Anatomy of an outbreak. Access at <https://www.facebook.com/siddharth.sridhar.5/posts/10158917399409258>
- Tosepu, R., Gunawan, J., Effendy, D. S., Ahmad, L.O. A. I., Lestari, H., Bahar, H., &Asfian, P. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Science of the Total Environment*, 725 (2020) 138436.
- Tuckel, Peter, Sassler, Sharon L., Maisel, Richard, Leykam, Andrew, 2006. The diffusion of the influenza pandemic of 1918 in Hartford, Connecticut. *Soc. Sci. History* 30 (2), 167–196.
- Vuorinen, V., Aarnio, M., Alava, M., Alopaeus, V., Atanasova, N., Auvinen, M., ... & Österberg, M. (2020). Modelling aerosol transport and virus exposure with numerical simulations in relation to SARS-CoV-2 transmission by inhalation indoors. *Safety Science*, 130, 104866.
- Wang, M., Jiang, A., Gong, L., Lu, L., Guo, W., Li, C., ...Li, H.(2020). Temperature Significantly Change COVID-19 Transmission in 429 cities. *medRxiv preprint*.
- Wei, Ye, Song, Wei, Xiu, Chunliang, & Zhao, Ziyu. (2018). The rich-club phenomenon of China's population flow network during the country's spring festival. *Applied Geography*, 96, 77-85.
- World Health Organization (WHO). (2020). Key planning recommendations for mass gatherings in the context of COVID-19 (Interim guidance). <https://www.who.int/publications/i/item/10665-332235>
- Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*.
- Yang, Z., Zeng, Z., Wang, K., Wong, S., Liang, W., Zanin, M.,...& He, J. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease*, 12(3), 165-174.
- Yezli, S., & Khan, A. (2020). COVID-19 social distancing in the Kingdom of Saudi Arabia: Bold measures in the face of political, economic, social and religious challenges. *Travel Medicine and Infectious Disease*, 37, 101692.
- Ye D., Zhang Q., Dong W., Chen Z., &Zhao Z. (2004). Analysis on the relationship between meteorological conditions and SARS occurring. *Climatic and Environmental Research*, 9(4), 670- 679.
- Zhao, S., Zhuang, Z., Ran, J., Lin, J., Yang, G., Yang, L., & He, D. (2020). The association

between domestic train transportation and novel coronavirus outbreak in China, from 2019 to 2020: A data-driven correlational report. *Travel Medicine and Infectious Disease*, 101568.

Zhou, L., Li, S., Li, C., Shen, G., Yang, H., Zhu, P., ... & Li, B. (2022). Spatial congruency or mismatch? Analysing the COVID-19 potential infection risk and urban density as businesses reopen. *Cities*, 103615.

Zhu, P., & Tan, X. (2021). Is compulsory home quarantine less effective than centralized quarantine in controlling the COVID-19 outbreak? Evidence from Hong Kong. *Sustainable Cities and Society*, 74, 103222.

Zhu, P., & Guo, Y. (2021). The role of high-speed rail and air travel in the spread of COVID-19 in China. *Travel medicine and infectious disease*, 42, 102097.

Zhu, P., Li, J., & Hou, Y. (2022). Applying population-flow-based spatial weight matrix in spatial econometric models: conceptual framework and application to COVID-19 transmission analysis. *Annals of the American Association of Geographers*.

Appendix

Donor cities and weights for synthetic controls in the robustness tests

| Panel B | | Panel C | | Panel D | | Panel E | |
|--------------|--------|--------------|--------|-----------|--------|--------------|--------|
| City | Weight | City | Weight | City | Weight | City | Weight |
| Shanghai | 0.292 | Macau | 0.294 | Chengdu | 0.31 | Changdu | 0.25 |
| Macau | 0.265 | Beijing | 0.193 | Macau | 0.254 | Macau | 0.243 |
| Longnan | 0.124 | Longnan | 0.151 | Beijing | 0.206 | Beijing | 0.216 |
| Shuangyashan | 0.078 | Shuangyashan | 0.146 | Shanghai | 0.146 | Dazhou | 0.154 |
| Beijing | 0.075 | Dazhou | 0.1 | Changdu | 0.058 | Shuangyashan | 0.128 |
| Guigang | 0.059 | Guigang | 0.074 | Wuhan | 0.016 | Wuhan | 0.009 |
| Shantou | 0.051 | Shangrao | 0.023 | Suqian | 0.007 | | |
| Wuhan | 0.016 | Wuhan | 0.019 | Xiaogan | 0.005 | | |
| Changdu | 0.037 | | | | | | |
| Xiaogan | 0.003 | | | | | Wuhan | 0.015 |
| Panel F | | Panel G | | Panel H | | Panel I | |
| City | Weight | City | Weight | City | Weight | City | Weight |
| Macau | 0.262 | Chongqing | 0.319 | Changdu | 0.3 | Macau | 0.393 |
| Beijing | 0.216 | Macau | 0.29 | Macau | 0.297 | Shenzhen | 0.305 |
| Shanghai | 0.134 | Beijing | 0.24 | Beijing | 0.228 | Beijing | 0.229 |
| Shantou | 0.19 | Changdu | 0.151 | Chongqing | 0.134 | Shanghai | 0.041 |
| Shuangyashan | 0.073 | | | Xiaogan | 0.02 | Wuhan | 0.015 |
| Changdu | 0.056 | | | Suqian | 0.015 | Dongying | 0.012 |
| Chongqing | 0.046 | | | Wuhan | 0.007 | Nanjing | 0.004 |
| Wuhan | 0.016 | | | | | | |
| Chuzhou | 0.005 | | | | | | |
| Huanggang | 0.001 | | | | | | |
| Panel J | | Panel K | | | | | |
| City | Weight | City | Weight | | | | |
| Shanghai | 0.464 | Shenzhen | 0.373 | | | | |
| Macau | 0.282 | Macau | 0.337 | | | | |
| Changdu | 0.102 | Beijing | 0.23 | | | | |
| Tianjin | 0.055 | Shanghai | 0.042 | | | | |
| Haidong | 0.048 | Wuhan | 0.013 | | | | |
| Beijing | 0.022 | Nanjing | 0.005 | | | | |
| Wuhan | 0.014 | | | | | | |
| Shannan | 0.004 | | | | | | |
| Chongqing | 0.002 | | | | | | |