Quantifying the impact of the Tokyo Olympics on COVID-19 cases using synthetic control methods

Taro Esaka [a],[*], Takao Fujii [b]

[a] *Kobe City University of Foreign Studies, 9-1, Gakuenhigashi-machi, Nishi-ku, Kobe, Hyogo 651-2187, Japan*
[b] *Kobe City University of Foreign Studies, 9-1, Gakuenhigashi-machi, Nishi-ku, Kobe, Hyogo 651-2187, Japan and Research Fellow at the Graduate School of Economics, Kobe University, Japan*

ARTICLE INFO

ABSTRACT

This paper uses a synthetic control method (SCM) and a Ridge Augmented SCM to estimate the impact of holding the Tokyo Olympic games on the number of newly confirmed COVID-19 cases in Tokyo (Japan). Our analysis with these methods enables us to estimate the causal impact of the Tokyo Olympics on COVID-19 cases by constructing counterfactual COVID-19 cases for Tokyo (Japan) as the optimal weighted average of COVID-19 cases of OECD countries that are not affected by holding the Olympics through a data-driven approach. Based on reliable estimates obtained from different analytical settings, we find that, compared to the counterfactuals, holding the Tokyo Olympics significantly increased the daily average number of COVID-19 cases by 105 to 132 cases in Tokyo (47 to 65 cases in Japan as a whole) per million people. This result suggests that holding the Olympics likely led to the spread of COVID-19 infection in Tokyo (Japan).

1. Introduction

The Tokyo Olympic games, which were delayed for a year due to the global spread of COVID-19, were held from July 23 to August 8, 2021. Unlike normal Olympic games, all events were held without spectators in and around the host city of Tokyo because COVID-19 infection entered an expansion phase in Tokyo in early July. Whether the Olympics should have been held or cancelled while the COVID-19 pandemic was not yet under control was actively debated in many quarters, including the National Diet, academic societies, TV, newspapers, and social networking sites. In general, there was much skepticism about the safety of holding the Olympics. For example, according to a poll conducted by Asahi Shimbun (2021a) on July 17 and 18, 33% of respondents were in favor of holding the Olympics, while 55% were against. In the same survey, 21% of respondents said that a "safe and secure Olympics" could be achieved, while 68% said they could not.

Given the development of newly confirmed COVID-19 cases shown in Fig. 1, many people believe that the Olympics increased new COVID-19 cases in Tokyo because it appears that COVID-19 cases increased rapidly around the time of the Olympics. According to a poll conducted

by Mainichi Shimbun (2021) on August 28, more than 70% of respondents said that holding the Olympics might have affected the spread of COVID-19 infection. From a causal inference perspective, did holding the Tokyo Olympics truly increase the number of new COVID-19 cases in Tokyo? If so, how many cases did it cause? These are important questions of social concern and policy relevance. However, there have been no reliable post hoc studies using causal inference methods to answer these questions.

The aim of this paper is to answer these questions by estimating the causal impact of holding the Tokyo Olympics on newly confirmed COVID-19 cases in Tokyo (Japan) using a novel causal inference approach. To estimate this impact, we need to know the unobserved counterfactual COVID-19 cases; that is, what would have been the cases if the Olympics had not been held. To do so, we use a synthetic control method (SCM) proposed by Abadie et al. (2010, 2015) and a Ridge Augmented SCM (RASCM) developed by Ben-Michael et al. (2021). These methods allow us to estimate the causal impact of an event on its outcome for a single treatment unit by constructing a counterfactual so-called "synthetic control" (SC), which can be constructed from outcomes in the control units through a data-driven approach (Abadie et al.,
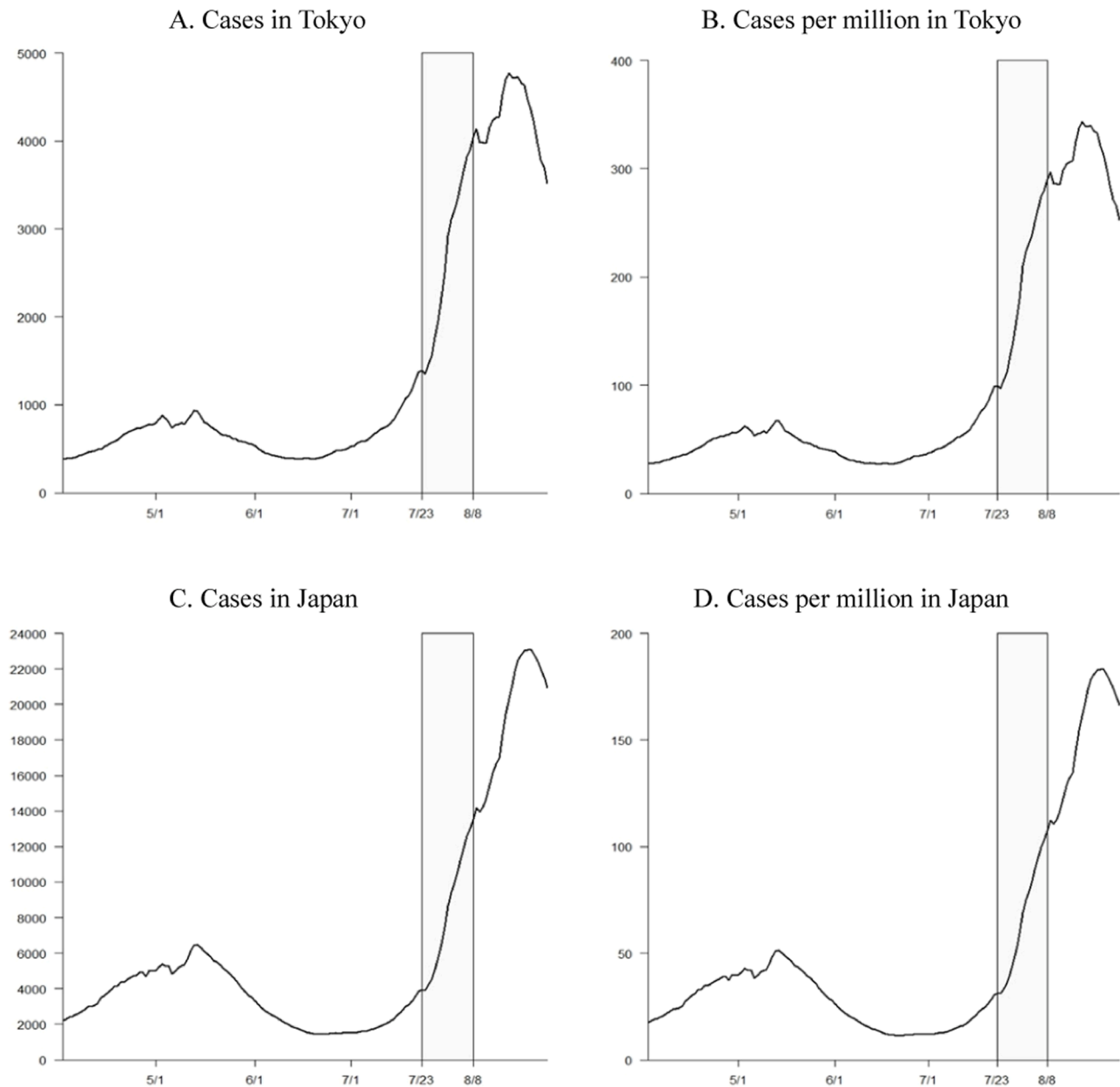
Fig. 1. The evolution of 7-day moving average of newly confirmed COVID-19 cases.
Notes: The shadow areas represent the duration of the Olympics.
Sources: Japanese Ministry of Health, Labour and Welfare and Our World in Data.
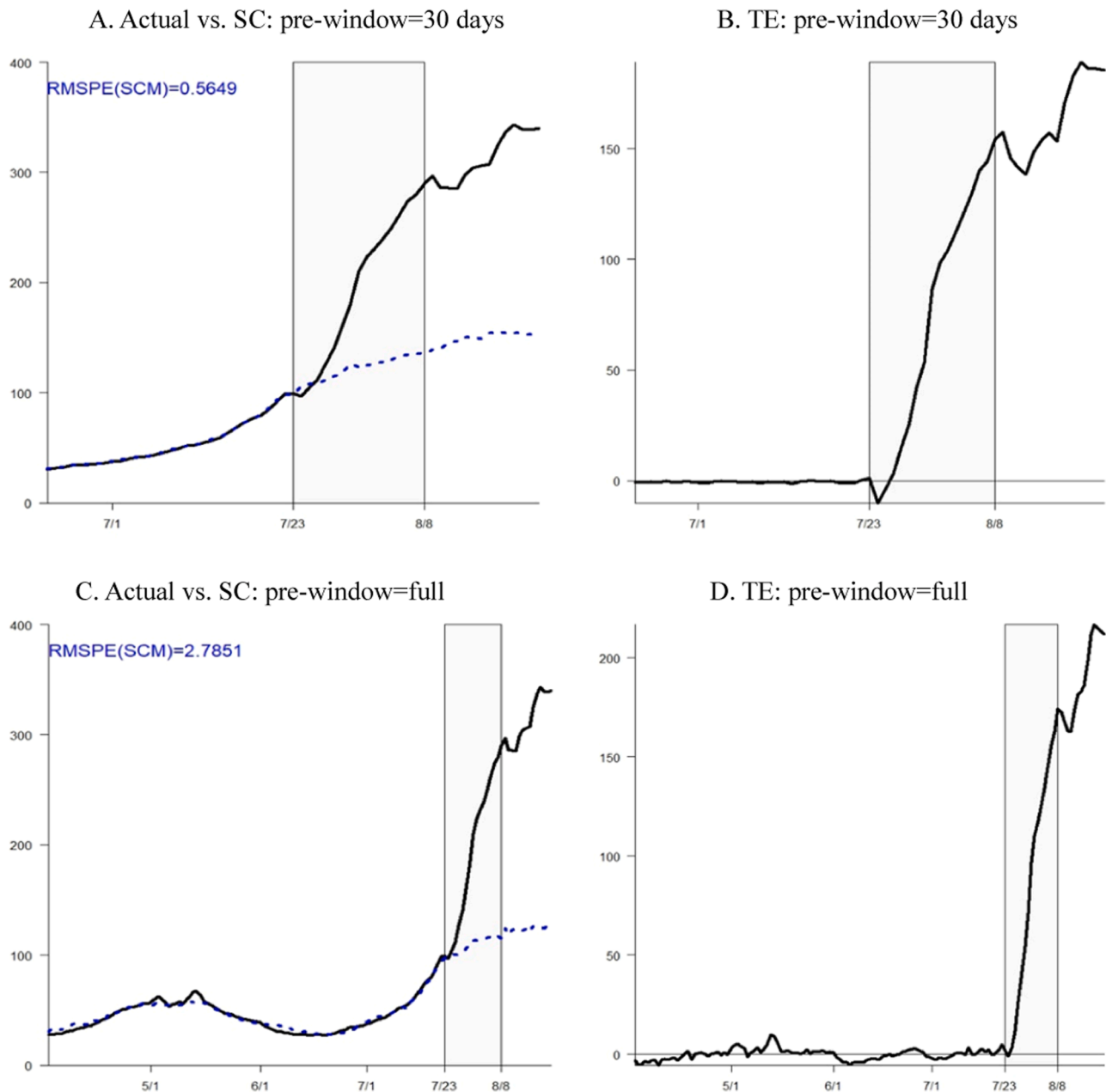
2010, 2015; Athey and Imbens, 2017; Abadie, 2021).

In our analysis using SCMs, we use a daily 7-day moving average of newly confirmed COVID-19 cases per one million people as an outcome variable and define holding the Olympics as a treatment (event), Tokyo (Japan) as a treatment unit, and 37 OECD countries as control units.[1] Following Doudchenko and Imbens (2016) and Botosaru and Ferman (2019), we use all lagged COVID-19 cases (i.e., all lagged outcomes) as predictors to construct an SC that can accurately fit the path of the actual outcome for the pre-treatment period. Using the SCMs, we construct a counterfactual for COVID-19 cases in Tokyo (Japan) as a weighted average of COVID-19 cases of the 37 OECD countries that are not affected by holding the Tokyo Olympics. We obtain a reliable counterfactual given that the estimated SC has a good pre-treatment fit in predicting the actual COVID-19 cases. According to the context of the SCM, using our reliable SC enables us to estimate the treatment effect of the Olympics by controlling for time-varying observed and unobserved factors that affect COVID-19 cases. The difference between the actual COVID-19 cases and their SC values in the post-treatment period indicates the treatment effect of holding the Olympics on COVID-19 cases in Tokyo (Japan).

We find that holding the Tokyo Olympics significantly increased the number of newly confirmed COVID-19 cases in Tokyo. This finding is qualitatively robust to varying pre-treatment windows, measurements of outcomes, and different SCMs. The quantitative impacts are slightly different for each analytical setting. When looking at the daily average of newly confirmed COVID-19 cases per one million people from July 23

---

[1] Although it would have been better to use Japanese prefectures without venues for the Olympics as control units (donors) in the case of the estimation for Tokyo, we did not adopt them as donors for many reasons, as shown in Section 3.2 and Appendix A.

A. Actual vs. SC: pre-window=30 days

RMSPE(SCM)=0.5649

B. TE: pre-window=30 days

C. Actual vs. SC: pre-window=full

RMSPE(SCM)=2.7851

D. TE: pre-window=full

Fig. 2. SCM estimates for Tokyo: level of the outcome.
Notes: Fig. 2 summarizes the SCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Tokyo when using the level of the number of the cases as the outcome. For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

to August 22, the Olympics increased COVID-19 cases by 105 to 132 in Tokyo (47 to 65 in Japan) relative to the counterfactuals. Based on these estimates and the population of Tokyo (Japan), our results suggest that if the Olympics had not been held, the daily average of new COVID-19 cases could have been reduced by as many as approximately 1500 to 1850 cases in Tokyo (approximately 5900 to 8150 cases in Japan as a whole). These results suggest that holding the Tokyo Olympics likely led to the spread of COVID-19 infection in Tokyo (Japan).

1.1. Related literature

This paper relates to three strands of the growing literature on the

COVID-19 pandemic. First, this paper is directly related to recent studies that examine the relationship between holding the Tokyo Olympics and COVID-19 infection rates. Professors Nakata and Fujii of the University of Tokyo build a traceable SIR-macro model and use it to conduct a simulation analysis of the relationship between economic output and the spread of COVID-19 in Japan (Fujii and Nakata, 2021).[2] As part of their

---

[2] They have performed a simulation analysis with new data every week to examine COVID-19 related policy issues, and updated the results on their website (https://covid19outputjapan.github.io/JP/). Their research project was the most successful in considering the relationship between the economy and COVID-19 in Japan in real time.

Table 1
SCM estimates for Tokyo.

RMSPE=0.5649	(1)	(2)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	113.25	3510.78
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(29.27, 197.23)	(907.47, 6114.10)
RMSPE=2.7851	(3)	(4)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	131.87	4087.88
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(34.09, 229.65)	(1056.64, 7119.12)
RMSPE=0.3417	(5)	(6)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	113.62	3522.36
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(29.37, 197.88)	(910.46, 6134.25)
RMSPE=5.7230	(7)	(8)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	141.81	4396.16
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(36.66, 246.97)	(1136.32, 7656.00)

Notes: From the SCM results when using the post-treatment period from July 23 to August 22, we calculate the average daily treatment effect (TE) and the cumulative TE in Tokyo. When using the log of COVID-19 cases as the outcome, the estimates obtained from the SCM are converted to the level of the outcomes (i.e., exp (SC)) to easily compare to the baseline results in (1) and (2). Then, we calculate the root mean squared prediction error (RMSPE), the average daily TE, and the cumulative TE. The two-sided p-values are calculated from the placebo and permutation tests that use the country-specific RMSPE ratio. The 90% confidence intervals are calculated from the two-sided p-values.

research, Fujii et al. (2021) present the results of model-based analyses on May 21 and June 16, 2021 to quantify the impact of holding the Olympics and Paralympics on COVID-19 infection in Tokyo. Based on June-16th analysis, they show that as of one week after the closing of the Olympics, the Olympics would increase the number of new COVID-19 cases in Tokyo by approximately 500 per day compared to the baseline of no Olympics. Fujii et al. (2021) point out that the festival moods promoted by holding the Olympics could lead to the spread of COVID-19 infection. Linton et al. (2021) employ a multitype branching process model by using information related to COVID-19 and the Olympics just before the Olympic Games to simulate the number of new COVID-19 cases in Tokyo after the opening of the Olympics. They show that if COVID-19 transmission could not be controlled, the maximum number of new cases in Tokyo would reach up to more than 4000 per day. Fujii et al. (2021) and Linton et al. (2021) perform ex ante simulation studies on the potential infection in Tokyo.

However, to the best of our knowledge, as of December 1, 2021, there have been no reliable ex post studies using causal inference methods, although more research will be conducted in the future. The daily average effect of holding the Olympics estimated from the SCMs is approximately 1500 to 1850 cases in Tokyo, indicating that our

estimates are different and larger than that of Fujii et al. (2021).[3] The reason for this is, of course, the difference in methodology, plus the fact that the study by Fujii et al. (2021) could not use the information that the number of new COVID-19 cases in Tokyo increased considerably in July.[4] Therefore, our contribution to this literature is to provide policy evaluation findings obtained from a reliable causal inference approach for estimating the impact of holding the Olympics on new COVID-19 cases in Tokyo (Japan).

Second, this paper is connected to empirical studies that examine the relation between sports events and COVID-19 infection in the local population. From a public health perspective, it is clear that sports events with many spectators substantially increase COVID-19 infection risk among spectators during the COVID-19 pandemic (e.g., Parnell et al., 2020). Therefore, do these events increase the risk of COVID-19 infection in the local population where they are held? To answer this question, Breidenbach and Mitze (2022) study the impact of hosting the German professional football matches, Toumi et al. (2021) study the impacts of hosting the United States National Football League games and NCAA football games, Carlin et al. (2021) study the impacts of hosting the NHL hockey games, NBA basketball games, and NCAA basketball games, and Dave et al. (2021) study the impact of the Sturgis Motorcycle Rally in South Dakota. Breidenbach and Mitze (2022), Carlin et al. (2021), and Dave et al. (2021) find that these events with large gatherings of people significantly increased the number of new COVID-19 cases in the local population, while Toumi et al. (2021) find no significant evidence. The first main difference between these studies and ours is that we examine the impact of holding a national event, such as the Olympics, on COVID-19 infection in the local population. The second main difference is that we estimate the impact of holding a large-scale event without spectators on COVID-19 infection in the local population. Therefore, our contribution is to provide empirical evidence that holding the Tokyo Olympics significantly increased the number of new COVID-19 cases in the local population, even though there was no audience.

Third, our paper is related to recent COVID-19 studies in application of the SCM. Using the SCM, Friedson et al. (2021) conduct a policy evaluation of shelter-in-place orders in California, Mitze et al. (2020) estimate the preventive effect of face masks in Germany, Cho (2020) examine the effectiveness of lockdown by estimating a counterfactual case for Sweden, Dave et al. (2021) estimate the impact of the Sturgis Motorcycle Rally in South Dakota on COVID-19 infection, and Breidenbach and Mitze (2022) estimate the impact of hosting professional football games on COVID-19 infection in Germany. Although these studies use the standard SCM proposed by Abadie et al. (2010, 2015), our study uses a RASCM developed by Ben-Michael et al. (2021) in addition to the standard SCM to estimate the impact of holding the Tokyo Olympics on new COVID-19 cases in Tokyo. The RASCM can construct an SC that better fits the trajectory of the actual outcome in the pre-treatment period than the standard SCM by admitting small negative weights. Using the RASCM, we obtain a more reliable counterfactual in the sense that the estimated SC has a better pre-treatment fit. Therefore, by comparing the SCM estimates to the RASCM estimates, we can increase the reliability of our results for COVID-19 infection.

---

[3] The effect of holding the Olympics was not certain in Linton et al. (2021) because it was not compared to the case where if the Olympics were not held. Thus, it is not possible to simply compare Linton et al. (2021) estimates to ours.

[4] It is natural that their May-21st and June-16th analyses could not use COVID-19-related information for future July. The study by Fujii et al. (2021) is important and significant because as of two months before the Olympics, it presented that holding the Olympics would greatly increase the spread of COVID-19 in Tokyo.
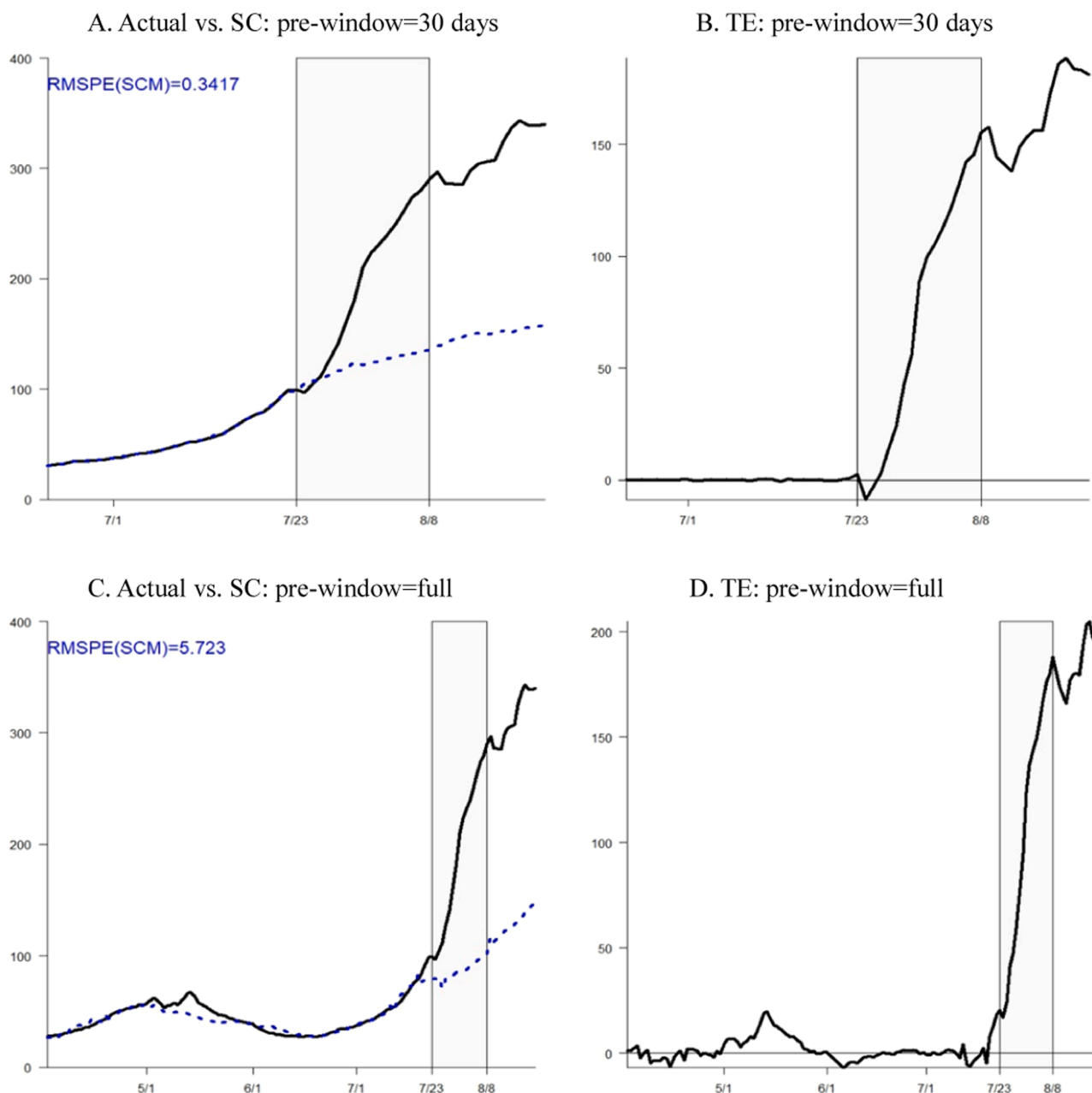
Fig. 3. SCM estimates for Tokyo: log of the outcome.

Notes: Fig. 3 summarizes the SCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Tokyo when using the log of the number of the cases as the outcome. To easily compare to the level of the outcome in Fig. 2, the estimates obtained from the SCM are converted to the level of the outcomes (i.e., exp (SC)). For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.
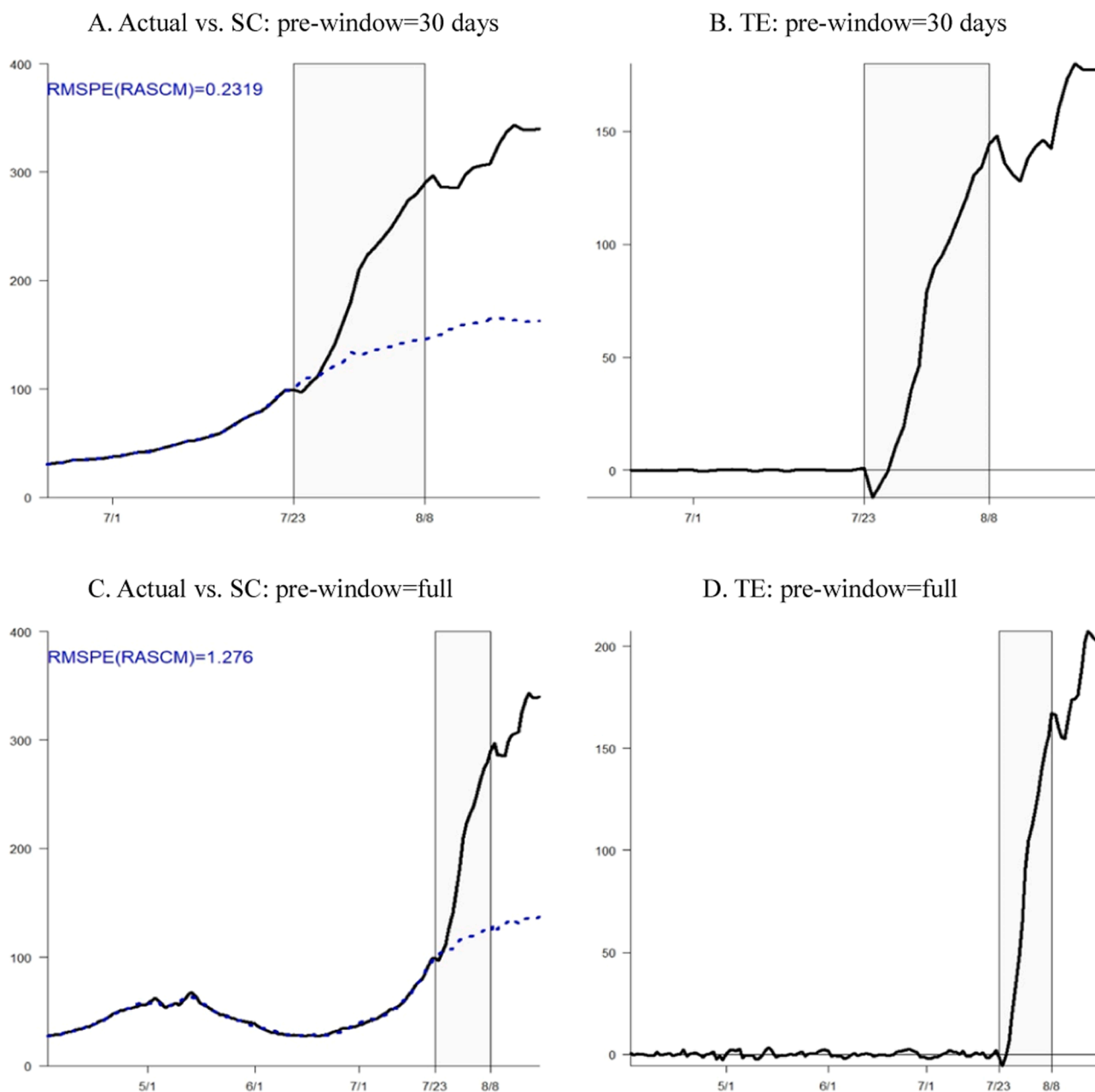
2. Background

2.1. COVID-19 cases in Tokyo before the Olympics and control measures

Tokyo is the capital and the economic and administrative center of Japan. The population was approximately 14 million in 2020, and the GDP was approximately 107 trillion yen (965.4 billion dollars) in fiscal 2018.[5] Comparing Tokyo to the 38 OECD countries, Tokyo ranked 17th in population and 12th in GDP. Hence, Tokyo is a super megacity in Japan that is comparable to the major countries of the world.

Tokyo and Japan had a relatively lower number of newly confirmed COVID-19 cases than OECD countries. As shown in Fig. 1, in early April 2021, the number of new COVID-19 cases per day was approximately 400 (30 per million people) in Tokyo, while it was approximately 2000 (20 per million people) in Japan as a whole. However, new COVID-19 cases increased until mid-May, reaching approximately 900 (70 per million people) in Tokyo and approximately 6000 (50 per million

---

[5] See Tokyo Statistics (https://www.toukei.metro.tokyo.lg.jp/).

A. Actual vs. SC: pre-window=30 days

RMSPE(RASCM)=0.2319

B. TE: pre-window=30 days

C. Actual vs. SC: pre-window=full

RMSPE(RASCM)=1.276

D. TE: pre-window=full

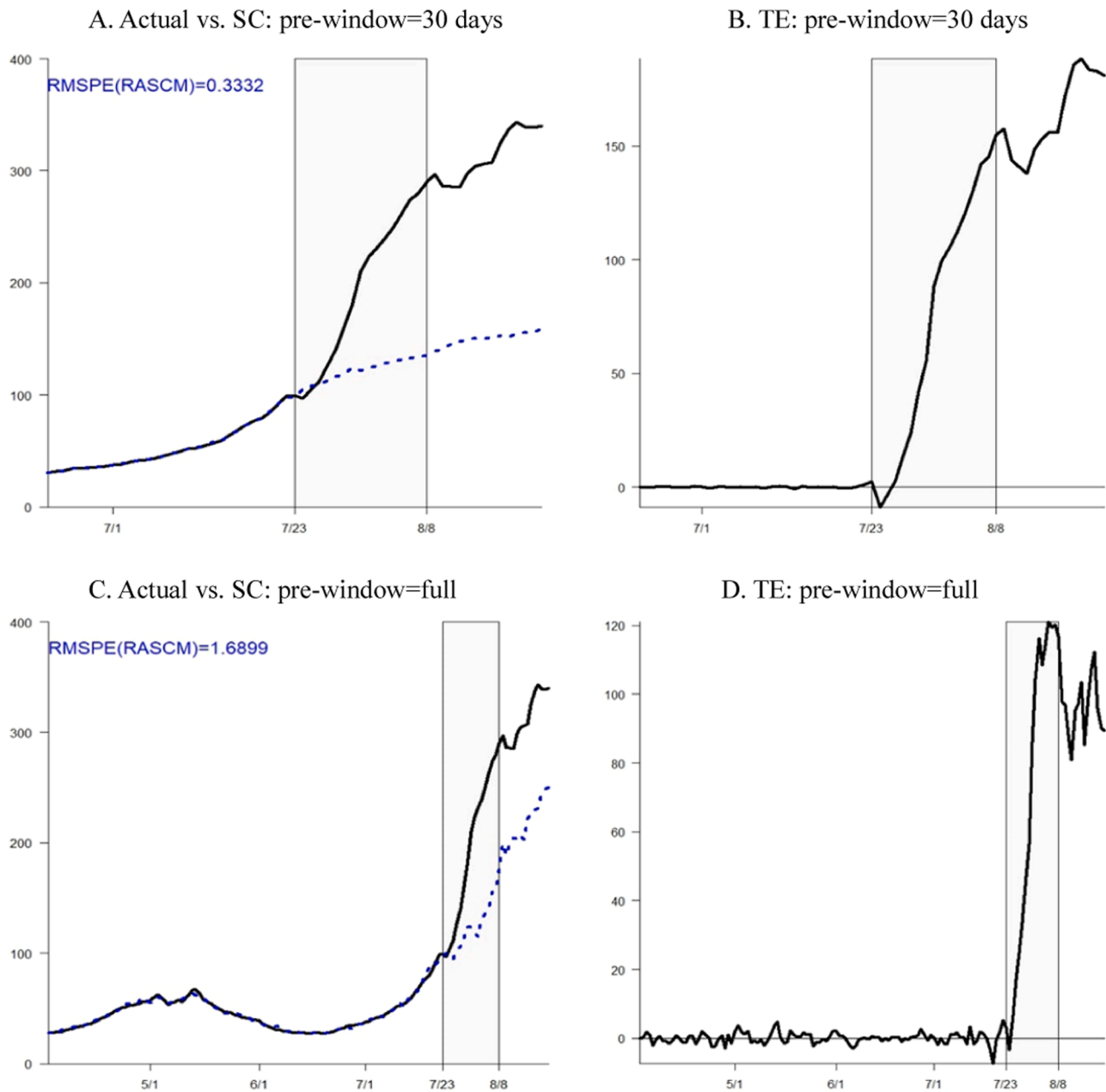**Fig. 4.** RASCM estimates for Tokyo: level of the outcome.
Notes: Fig. 4 summarizes the RASCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Tokyo when using the level of the number of the cases as the outcome. For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

people) in Japan. Since then, the number of new cases had been decreasing, and by mid-June, it had dropped to the level of early April. In response to this decrease, the Japanese government lifted the third state of emergency that was in place in Tokyo from April 25 to June 20.[6]

In early July, the number of new COVID-19 cases began to increase again. In response to the spread of the infection, on the night of July 8, the government declared the fourth state of emergency in Tokyo from July 12 to August 22. At the same time, the Tokyo Organising Committee of the Olympic and Paralympic Games (TOCOG) announced that all events for the Olympics would be held without spectators in and around the host city of Tokyo to hold the Olympics while preventing the spread of COVID-19 infection. From a public health perspective, holding the Olympics was inconsistent with the state of emergency for preventing the spread of COVID-19. Never before have the Olympics been held during a new virus spread as was the case with the Tokyo Olympics. Therefore, our research is a case study to estimate the impact of a large-scale and rare event, such as the Olympics, on new viral infections, such

---

[6] Although many countries imposed a lockdown with restrictions to prevent COVID-19 infection, the state of emergency government in Japan was not enforceable and nonpunitive with the declaration.

A. Actual vs. SC: pre-window=30 days

B. TE: pre-window=30 days

C. Actual vs. SC: pre-window=full

D. TE: pre-window=full

Fig. 5. RASCM estimates for Tokyo: log of the outcome.
Notes: Fig. 5 summarizes the RASCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Tokyo when using the log of the number of the cases as the outcome. To easily compare to the level of the outcome in Fig. 4, the estimates obtained from the RASCM are converted to the level of the outcomes (i.e., exp (SC)). For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

as COVID-19, in the host city of Tokyo.

In addition to holding the events without spectators, the TOCOG implemented several measures to prevent the spread of COVID-19 infection.[7] In particular, the TOCOG adopted a "bubble system" in which athletes and participants were isolated from the outside world. Moreover, the committee made it mandatory for Olympic participants to follow the "Playbooks," a set of guidelines for infection control.

2.2. COVID-19 cases in Tokyo after the Olympics

The Tokyo Olympics were held from July 23 to August 8, 2021 under the state of emergency during the COVID-19 pandemic. The Olympics featured 339 events in 33 sports and attracted 11,259 athletes, approximately 33,000 overseas game officials, and 51,672 volunteers, among Olympic participants.[8] In addition to Tokyo, events were held in eight prefectures including Saitama, Kanagawa, Chiba, Ibaraki, Shizuoka, Fukushima, Miyagi, and Hokkaido. Almost all of the competition

---

[7] For details of COVID-19 countermeasures and Playbooks, see https://www.2020games.metro.tokyo.lg.jp/special/eng/guide/ and https://olympics.com/ioc/tokyo-2020-playbooks.

[8] See TOCOG (2021a).

Table 2
RASCM estimates for Tokyo.

RMSPE=0.2319	(1)	(2)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	105.18	3260.69
Placebo test: rank	2/38	2/38
Two-sided p-value	[0.0526]	[0.0526]
90% confidence interval	(15.76, 194.61)	(488.47, 6032.91)
RMSPE=1.2760	(3)	(4)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	124.86	3870.66
Placebo test: rank	3/38	3/38
Two-sided p-value	[0.0789]	[0.0789]
90% confidence interval	(7.81, 241.91)	(242.14, 7499.18)
RMSPE=0.3332	(5)	(6)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	113.46	3517.35
Placebo test: rank	2/38	2/38
Two-sided p-value	[0.0526]	[0.0526]
90% confidence interval	(17.00, 209.93)	(526.92, 6507.77)
RMSPE=1.6899	(7)	(8)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	81.65	2531.00
Placebo test: rank	6/38	6/38
Two-sided p-value	[0.1579]	[0.1579]
90% confidence interval	(-13.33, 176.62)	(-413.13, 5475.13)

Notes: From the RASCM results when using the post-treatment period from July 23 to August 22, we calculate the average daily treatment effect (TE) and the cumulative TE in Tokyo. When using the log of COVID-19 cases as the outcome, the estimates obtained from the RASCM are converted to the level of the outcomes (i.e., exp (SC)) to easily compare to the baseline results. Then, we calculate the root mean squared prediction error (RMSPE), the average daily TE, and the cumulative TE. The two-sided p-values are calculated from the placebo and permutation tests that use the country-specific RMSPE ratio. The 90% confidence intervals are calculated from the two-sided p-values.

Table 3
SCM estimates for Japan.

RMSPE=0.3495	(1)	(2)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	52.75	1635.22
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(13.63, 91.86)	(422.67, 2847.77)
RMSPE= 3.0003	(3)	(4)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	64.67	2004.71
Placebo test: rank	3/38	3/38
Two-sided p-value	[0.0789]	[0.0789]
90% confidence interval	(4.05, 125.29)	(125.41, 3884.01)
RMSPE=0.0912	(5)	(6)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	47.30	1466.20
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(12.23, 82.37)	(378.99, 2553.42)
RMSPE=6.0802	(7)	(8)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	37.89	1174.74
Placebo test: rank	9/38	9/38
Two-sided p-value	[0.2368]	[0.2368]
90% confidence interval	(-14.54, 90.33)	(-450.71, 2800.19)

Notes: From the SCM results when using the post-treatment period from July 23 to August 22, we calculate the average daily treatment effect (TE) and the cumulative TE in Japan. When using the log of COVID-19 cases as the outcome, the estimates obtained from the SCM are converted to the level of the outcomes (i.e., exp (SC)) to easily compare to the baseline results in (1) and (2). Then, we calculate the root mean squared prediction error (RMSPE), the average daily TE, and the cumulative TE. The two-sided p-values are calculated from the placebo and permutation tests that use the country-specific RMSPE ratio. The 90% confidence intervals are calculated from the two-sided p-values.

venues were in Tokyo and other metropolitan areas, and events took place without spectators.

Although the TOCOG implemented various infection control measures, how many of the Olympic participants were actually infected with COVID-19? The report by Japan Broadcasting Corporation (NHK) on September 8 indicates that the number of people infected with COVID-19 directly related to the Olympics was 547, which was based on an announcement by the TOCOG.[9] Of these 547 COVID-19 cases, 28 were athletes who came to Japan from overseas, 147 were from the IOC or sports organizations or were coaches and other officials, 32 were members of the media, 15 were from the organizing committee, 296 were contractors, and 29 were volunteers. Based on these figures and the low positivity rate of the screening test for COVID-19, the TOCOG (2021b) claimed that COVID-19 countermeasures were effective in controlling the spread of COVID-19 among Olympic participants.

What happened to COVID-19 cases among the citizens of the host city, Tokyo, as a result of holding the Tokyo Olympics? As shown in Fig. 1, it appears that the number of new COVID-19 cases in Tokyo

increased during the Olympics. Although the number of COVID-19 cases in Tokyo per day was approximately 1400 (100 per million people) on July 22, it was approximately 4000 (300 per million people) on August 8. In response to the increase in COVID-19 cases, the government announced on July 31 that the fourth state of emergency in Tokyo would be extended until August 31. On August 17, the government announced that it would be extended again until September 12. Finally, the fourth state of emergency was extended until September 30. The data in Fig. 1 lead many people to believe that the Olympics had caused the spread of COVID-19 in Tokyo.[10] However, the TOCOG (2021b) rejected this public opinion, citing the reasons that the effective reproduction number of COVID-19 had been on a downward trend and that people's mobility had been declining during the Olympics.

Unlike the TOCOG (2021b), using a causal inference approach, we address the question of whether holding the Olympics substantially increased the number of new COVID-19 cases in Tokyo. Specifically, we

---

[9] https://www3.nhk.or.jp/news/html/20210908/k10013249701000.html.

[10] For example, according to a poll conducted by Kyodo News (2021) on August 16, 59.8% of respondents thought that the holding the Olympics was a factor in the spread of COVID-19 infection.

Table 4

RASCM estimates for Japan.

RMSPE=0.0848	(1)	(2)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	50.83	1575.64
Placebo test: rank	2/38	2/38
Two-sided p-value	[0.0526]	[0.0526]
90% confidence interval	(7.61, 94.04)	(236.04, 2915.23)
RMSPE=1.1893	(3)	(4)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
	Average daily TE	Cumulative TE
TE	44.55	1380.92
Placebo test: rank	13/38	13/38
Two-sided p-value	[0.3421]	[0.3421]
90% confidence interval	(-31.74, 120.83)	(-983.83, 3745.68)
RMSPE=0.0898	(5)	(6)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	47.37	1468.61
Placebo test: rank	1/38	1/38
Two-sided p-value	[0.0263]	[0.0263]
90% confidence interval	(12.25, 82.50)	(379.61, 2557.61)
RMSPE=0.9598	(7)	(8)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Log (cases)	Log (cases)
	Average daily TE	Cumulative TE
TE	17.85	553.26
Placebo test: rank	8/38	8/38
Two-sided p-value	[0.2105]	[0.2105]
90% confidence interval	(-5.51, 41.21)	(-170.90, 1277.42)

Notes: From the RASCM results when using the post-treatment period from July 23 to August 22, we calculate the average daily treatment effect (TE) and the cumulative TE in Japan. When using the log of COVID-19 cases as the outcome, the estimates obtained from the RASCM are converted to the level of the outcomes (i.e., exp (SC)) to easily compare to the baseline results. Then, we calculate the root mean squared prediction error (RMSPE), the average daily TE, and the cumulative TE. The two-sided p-values are calculated from the placebo and permutation tests that use the country-specific RMSPE ratio. The 90% confidence intervals are calculated from the two-sided p-values.

use the SCM approach to estimate the counterfactual COVID-19 cases; that is, what would have been the cases if the Olympics had not been held. We then estimate the impact of holding the Olympics on COVID-19 cases in Tokyo by comparing the actual cases to the counterfactual cases. Therefore, our analysis does not examine the effectiveness of various COVID-19 countermeasures implemented in response to the Olympics but rather the overall impact of holding the Olympics on COVID-19 cases in Tokyo.

3. Methodology and data

To estimate the impact of holding the Tokyo Olympics on COVID-19 cases in Tokyo, we use the SCM proposed by Abadie et al. (2010, 2015) and the RASCM developed by Ben-Michael et al. (2021) which is an extended version of the SCM. Since the SCM was introduced in the seminal papers of Abadie and Gardeazabal (2003) and Abadie et al. (2010), it has been widely used to estimate the effect of a treatment for a single treatment unit in comparative case studies in various fields (Athey and Imbens, 2017; Bouttell et al., 2018; Abadie, 2021).

3.1. Synthetic control method

In this section, we present an empirical methodology of using the SCM. Following the treatment effect literature, we define holding the Tokyo Olympics as "treatment," Tokyo (Japan) as the "treatment unit," and countries other than Tokyo (Japan) as the "control unit." In our application of the SCM, we also define a daily 7-day moving average of newly confirmed COVID-19 cases per one million people as the "outcome" variable (Y_{it}). Y_{it} is the outcome in country i at time t.

We begin by presenting some notations of the SCM.[11] Suppose that we observe $J+1$ units over the period from time 1 to time T ($t = 1, \cdots, T_0, T_0 + 1 \cdots, T$). Without loss of generality, suppose that only the first unit is exposed to the treatment of interest (i.e., Tokyo is affected by the Tokyo Olympics), so that J remaining units are defined as potential control units (i.e., donor pool). Let Y_{1t} denote the observed outcome for Tokyo at time t and Tokyo be exposed to the treatment at time $T_0 + 1$. Let T_0 denote the number of pre-treatment periods, with $1 \leq T_0 \leq T$. Let Y_{it}^N be the counterfactual outcome for unit i at time t (i.e., what would have been an outcome if a treatment had not been adopted).

The SCM assumes that Y_{it}^N follows a factor model given by

$$Y_{it}^N = \eta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}, \tag{1}$$

where η_t is an unknown common factor (e.g., time-specific effect), θ_t is a $(1 \times r)$ vector of unknown parameters, Z_i is an $(r \times 1)$ vector of observed covariates, λ_t is a $(1 \times F)$ vector of unobserved common factors that depends on time, μ_i is an $(F \times 1)$ vector of unknown factor loadings (e. g., country-specific effect), and ε_{it} is unobserved transitory shock at a unit level with zero mean.

We aim to estimate the treatment effect (TE) of the treatment as follows:

$$\alpha_{1t} = Y_{1t} - Y_{1t}^N, \tag{2}$$

for $t > T_0$. However, we cannot directly estimate the TE (α_{1t}) because Y_{1t}^N cannot be observed. According to Abadie et al. (2010, 2015), the counterfactual outcome Y_{1t}^N for the treatment unit is called synthetic control (SC).

Consider a $(J \times 1)$ vector of weights $\mathbf{W} = (w_2, \cdots, w_{J+1})'$ such that $w_i \geq 0$ for $j = 2, \cdots, J+1$ and $w_2 + \cdots + w_{J+1} = 1$. In the SCM, Y_{1t}^N can be constructed from a weighted average of outcomes in the donor pool (i.e., $\sum_{j=2}^{J+1} w_j^* Y_{jt}$) through a data-driven approach. Abadie et al. (2010, 2015) argue that retaining the simplex constraint on the weights can avoid bias resulted from extrapolating outside the convex hull of the control units, indicating that the SCM prevents estimation of extreme counterfactuals (Ben-Michael et al., 2021). Therefore, the estimated TE is given by

$$\widehat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}, \tag{3}$$

for $t > T_0$. Abadie et al. (2010, 2015) show that under regular conditions, given optimal weights w_j^*, the estimated TE ($\widehat{\alpha}_{1t}$) is an unbiased estimator.

In our analysis, we need to choose the optimal weight w_j^* so that the SC can accurately fit the trajectory of the actual outcome before the start of the Tokyo Olympics because, theoretically, the relation $Y_{1t} = Y_{1t}^N$ holds for the pre-treatment period. Thus, Abadie et al. (2010, 2015) propose that the optimal weight vector W^* is chosen to minimize the following distance between X_1 and $X_0 W$:

---

[11] The presentation is partly based on the theory and application of the SCM by Abadie et al. (2010, 2015), Doudchenko and Imbens (2016), and Ben-Michael et al. (2021).
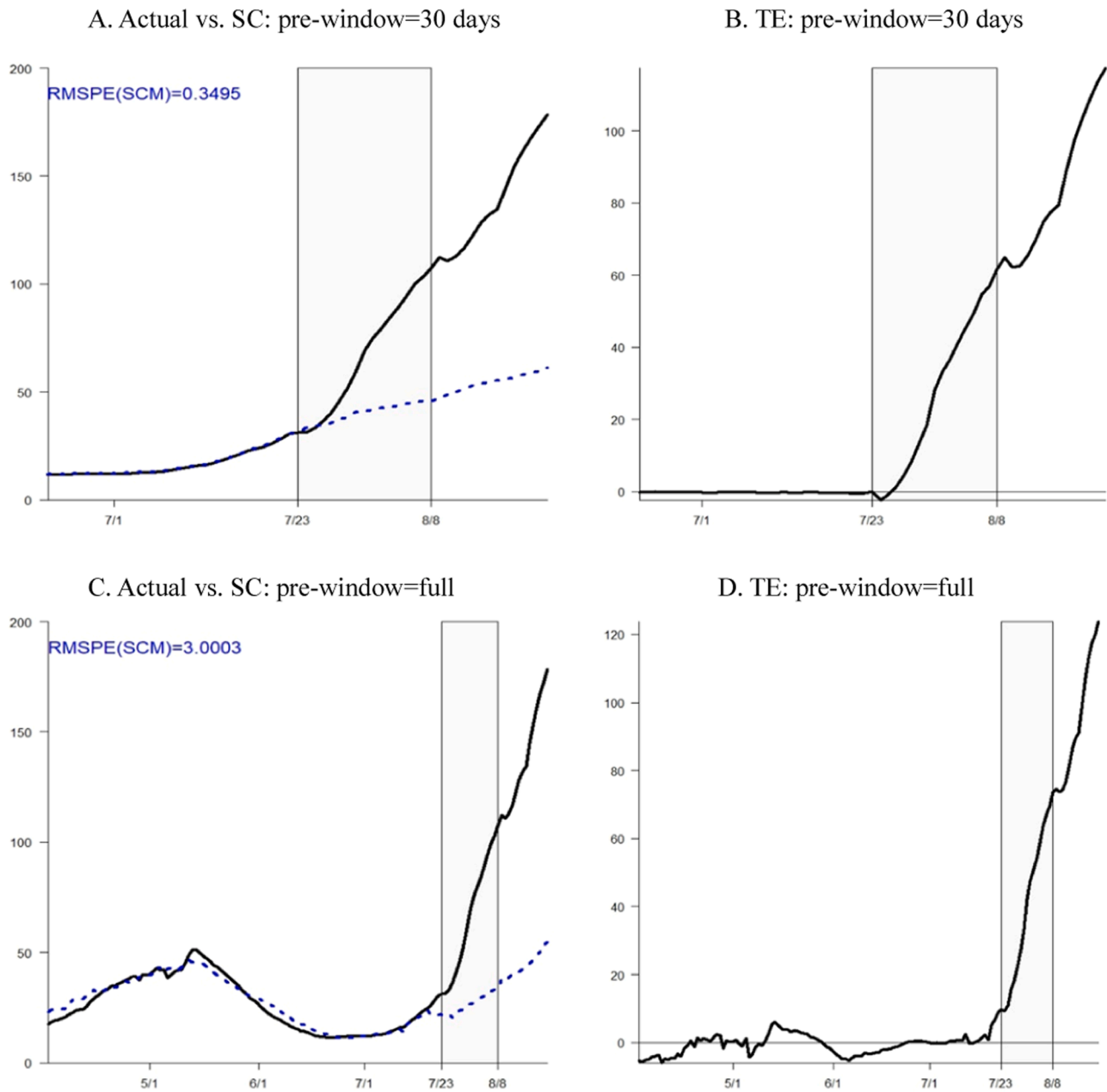
Fig. 6. SCM estimates for Japan: level of the outcome.

Notes: Fig. 6 summarizes the SCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Japan when using the level of the number of the cases as the outcome. For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

$$\| X_1 - X_0 W \|_v = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)},$$

where X_1 is a $(k \times 1)$ vector of pre-treatment predictors for the treated unit and, similarly, X_0 is a $(k \times J)$ matrix for the donor pool. X_1 and X_0 contain the pre-treatment outcomes Y and covariates Z. V is some $(k \times k)$ symmetric and positive semi-definite matrix that captures the relative importance of these variables as predictors of the outcome variable. The value of V is chosen such that the mean squared prediction error (MSPE) of the outcome variable is minimized for the pre-treatment period. To assess whether the estimated SC accurately fits the path of the actual outcome for the treated unit in the pre-treatment period, we calculate the root mean squared prediction error (RMSPE) of the SC and perform a

visual inspection.

According to Abadie et al. (2010, 2015), Abadie (2021), Bouttell et al. (2018), and Mitze et al. (2020), the assumptions and requirements for effective use of the SCM are as follows: (i) All control units do not adopt the treatment during the analysis period. (ii) The pre-treatment period is long (i.e., the number of pre-treatment observations is large). (iii) The treatment unit lies in the convex hull of control units. The outcomes and predictors for the treatment unit must not be extreme relative to those for the control units. (iv) There are no anticipation effects. In other words, there are no effects related to the treatment in the treatment unit before the event start date for the SCM. (v) In the use of predictors, covariates are not affected by the treatment. (vi) As in conventional treatment effect analyses, outcomes of the control units are
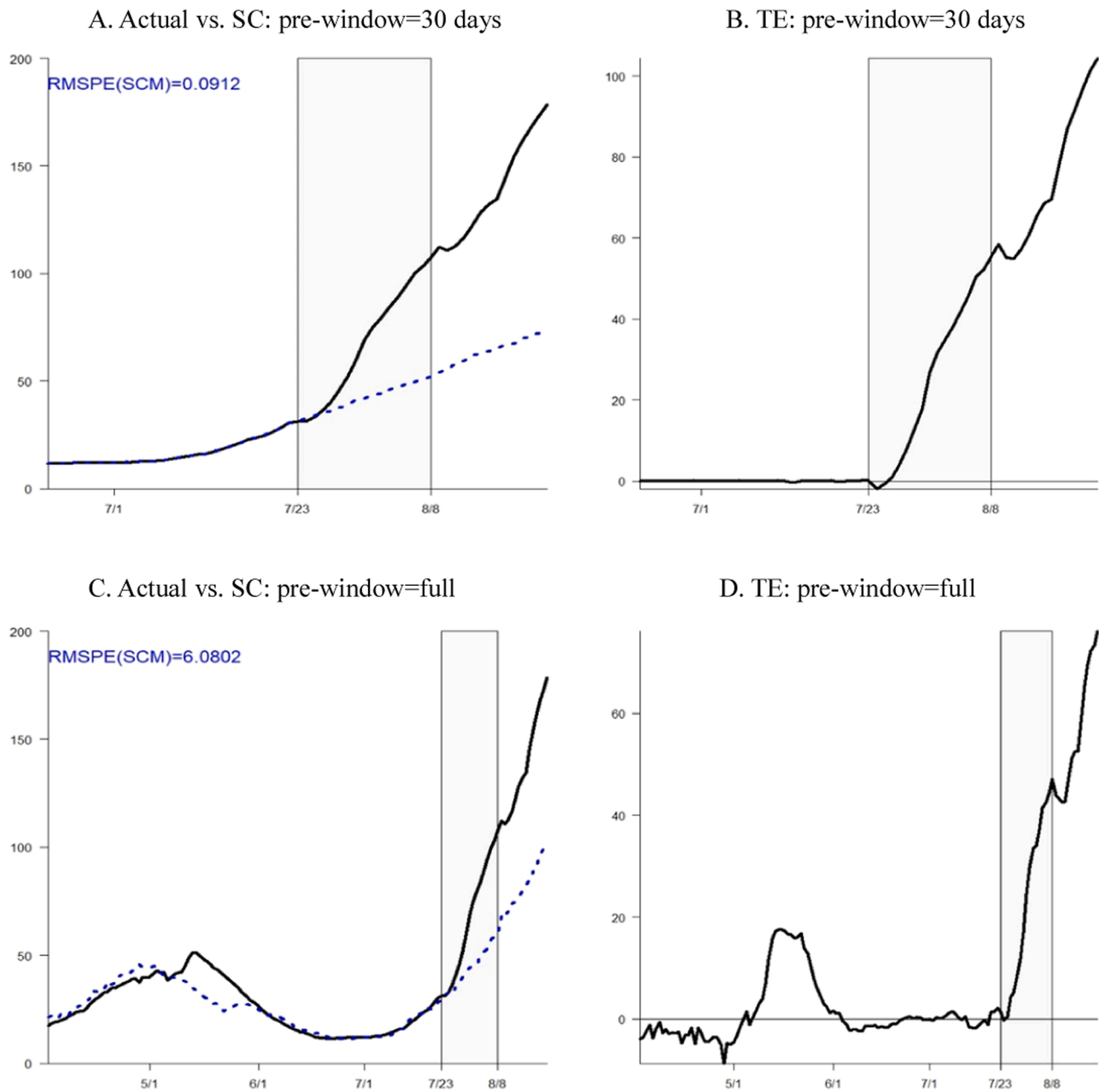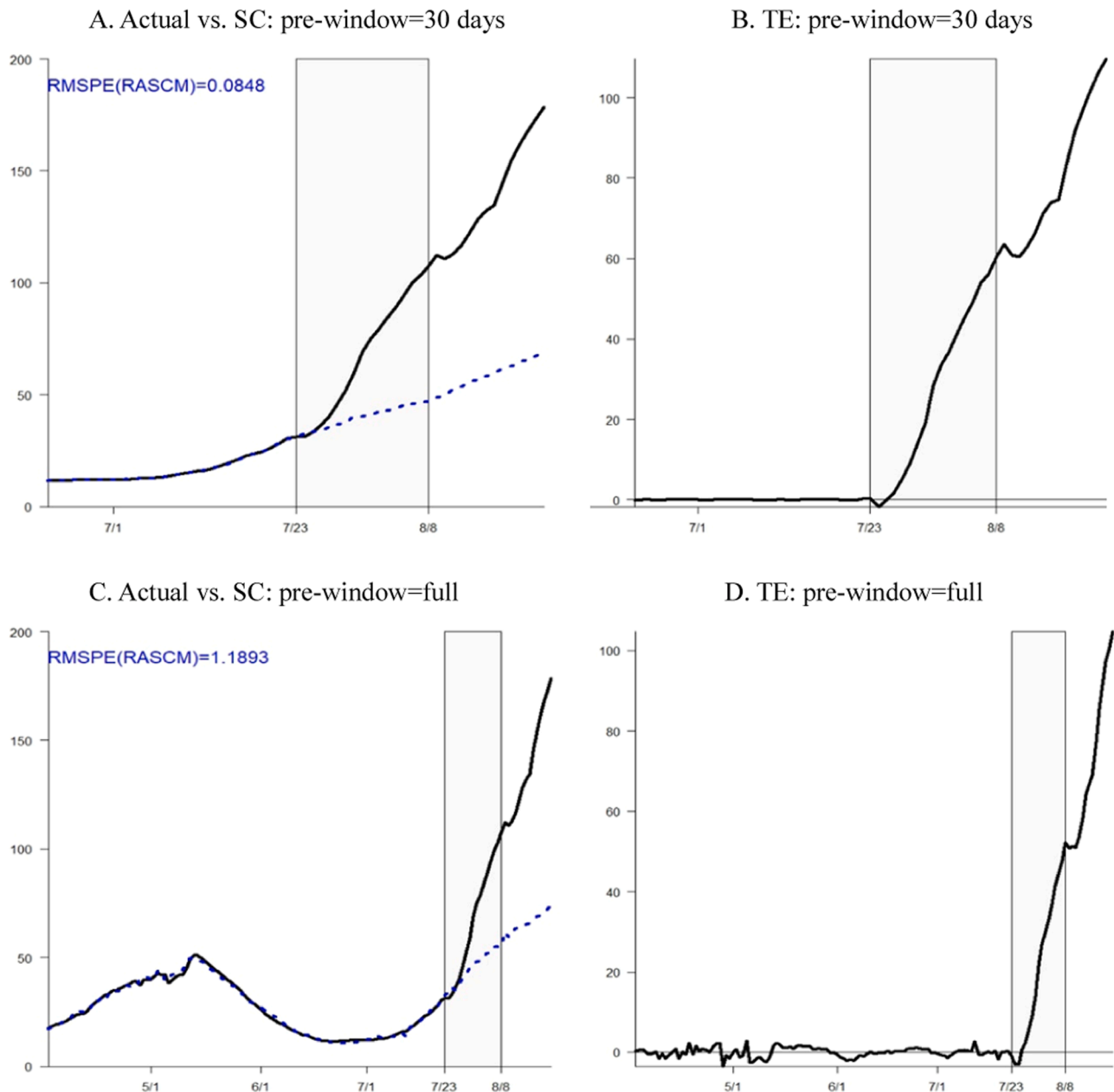
Fig. 7. SCM estimates for Japan: log of the outcome.
Notes: Fig. 7 summarizes the SCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Japan when using the log of the number of the cases as the outcome. To easily compare to the level of the outcome in Fig. 6, the estimates obtained from the SCM are converted to the level of the outcomes (i.e., exp (SC)). For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

not affected by the treatment in the treated unit. This condition is well known as the "stable unit treatment value assumption" (SUTVA). (vii) If it is not possible to construct an SC that can accurately fit the path of the actual outcome for the pre-treatment period, the counterfactual is not reliable. Accordingly, taking these conditions into account, we will select control units, pre-treatment periods, and prediction variables for estimating a reliable counterfactual outcome.

3.2. Data

Our analysis uses daily panel data from Tokyo and the 38 OECD countries from April 2, 2021 to August 22, 2021. The outcome variable is the number of daily newly confirmed COVID-19 cases per one million people. We use the 7-day moving average of the outcome to correct for the day-of-the week effect observed in COVID-19 data. The data are

A. Actual vs. SC: pre-window=30 days

RMSPE(RASCM)=0.0848

B. TE: pre-window=30 days

C. Actual vs. SC: pre-window=full

RMSPE(RASCM)=1.1893

D. TE: pre-window=full

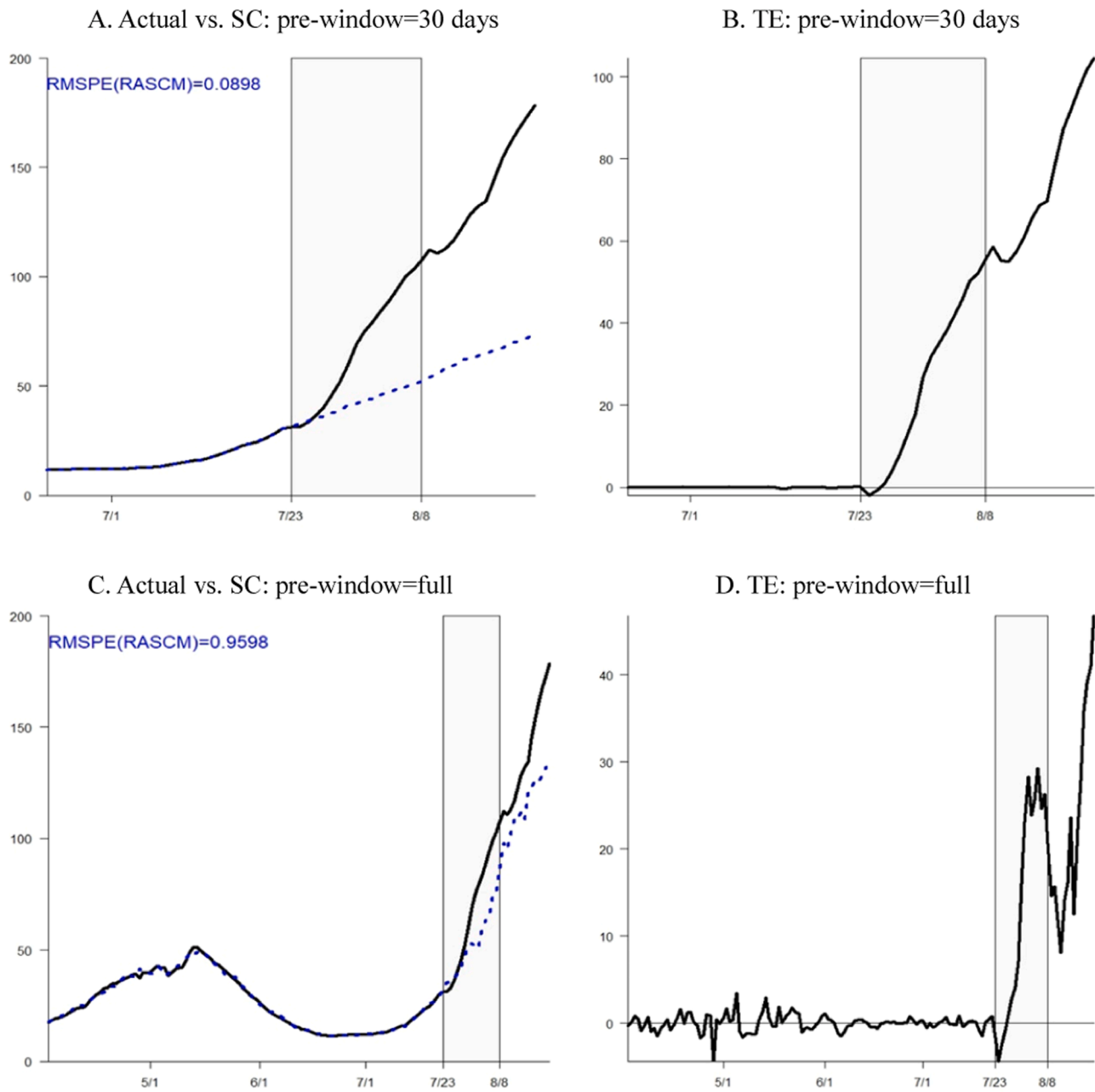**Fig. 8.** RASCM estimates for Japan: level of the outcome.
Notes: Fig. 8 summarizes the RASCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Japan when using the level of the number of the cases as the outcome. For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

obtained from the Our World in Data website (Ritchie et al., 2020).[12] Taking the SCM requirements into account, we select the 37 OECD countries except for Japan as our donor pool because the above-mentioned conditions (i), (iii), and (vi) are met. In fact, the number of COVID-19 cases in Tokyo was not large compared to those in other OECD countries, indicating that the treatment unit was in the convex hull of control units (see Appendix B). Importantly, it is unlikely that holding the Tokyo Olympics affected COVID-19 cases in the people in the 37 OECD countries, so we believe that the SUTVA is satisfied. In addition, despite being a city, Tokyo has an economy and population

comparable in size to those of an OECD country, and its social and political systems are similar to those of OECD countries.

Let us now consider the case if we use Japanese prefectures as the donor pool in the case of the estimation for Tokyo. As shown in Appendix A, the number of COVID-19 cases in Tokyo was much larger than that in Japanese prefectures, indicating that the treatment unit does not lie in the convex hull of control units. The SUTVA is not satisfied because it is possible that holding the Olympics had an impact on COVID-19 cases in other Japanese prefectures (i.e., there were spillover effects within Japan). Moreover, although we tried to use

Fig. 9. RASCM estimates for Japan: log of the outcome.

Notes: Fig. 9 summarizes the RASCM estimates from estimating the impact of the Tokyo Olympics on newly confirmed COVID-19 cases per million people in Japan when using the log of the number of the cases as the outcome. To easily compare to the level of the outcome in Fig. 8, the estimates obtained from the RASCM are converted to the level of the outcomes (i.e., exp (SC)). For using the pre-treatment period from June 23 to July 22, panels A and B show the daily actual COVID-19 cases and the SC (i.e., counterfactual) and the treatment effect (TE) of the Olympics, respectively. For using the pre-treatment period from April 2 to July 22, panels C and D show the daily actual COVID-19 cases and the SC and the TE of the Olympics, respectively. In panels A and C, the development of the actual COVID-19 cases is shown in the black solid line and the SC is shown in the blue dashed line. The shadow areas represent the duration of the Olympics.

Japanese prefectures as donors in the SCM, we could not construct a reliable counterfactual because the estimated SC did not fit the path of the actual outcome for the pre-treatment period.[13] Therefore, Japanese prefectures are not suitable for use as the donor pool for our analysis.

To construct an SC that can accurately fit the trajectory of the actual outcome for Tokyo in the pre-treatment period, we need to find reliable prediction variables. Abadie et al. (2010, 2015) show that if the SC can approximate the trajectory of the outcome for the treated unit in the pre-treatment period, it is possible to control for time-varying observed and unobserved factors that affect the outcome and the treatment. Doudchenko and Imbens (2016) show that lagged outcomes should be used as predictors because in terms of predictive power, they tend to be substantially more important, while in practice, other covariates tend to play a relatively minor role. Moreover, Abadie et al. (2010) suggest that researchers should not use covariates that are affected by the treatment. When following their suggestions, we cannot find appropriate covariates on a daily basis. Botosaru and Ferman (2019) show that if the SC obtained by using only the lagged outcomes as predictors has a perfect balance on pre-treatment outcome, it leads to an approximate balance for other covariates that may affect the outcome. This suggests that in application of the SCM, it would be possible to construct a reliable counterfactual even if there is no information on relevant covariates. Thus, following Doudchenko and Imbens (2016) and Botosaru and Ferman (2019), we use all lagged COVID-19 cases as predictors, indicating that X is equal to the full vector of pre-treatment outcomes Y_{it} for $t = 1, \cdots, T_0$ and contains no other variables.[14] Doudchenko and Imbens (2016) refer to this case of the SCM as constrained regression.

Our data period is from April 2, 2021 to August 22, 2021. Since the Tokyo Olympics were held from July 23 to August 8, 2021, we consider an analysis window consisting of 112 days before the start of the event, the event period (17 days), and 14 days after the closing of the event (i. e., 31 days after the start of the event). The number of observations before the event is large enough to apply the SCM. However, when using a long period including many periods of a low level of infection as pre-treatment window, the SC may be constructed from matching relatively small values of COVID-19 cases in the pre-treatment period. If so, the estimated counterfactual for Tokyo may be low, and the impact of holding the Olympics on COVID-19 cases may be excessively large. To address this concern, we use a 30-day pre-treatment window from June 23 to July 22 as a baseline window given that the SC may be more likely to match the growing path of COVID-19 cases just before the Olympics. The number of pre-treatment observations for our baseline is larger compared to recent COVID-19 studies in application of the SCM.[15] Emphasizing the length of pre-treatment window, we will use a full pre-treatment window from April 2 to July 22 as an alternative window. As noted by Friedson et al. (2021), when COVID-19 infection is growing at an exponential rate, it may be better to use the natural logarithm of the outcome rather than the level of the outcome. Considering such a condition, we will use the log of COVID-19 cases as an alternative outcome.

3.3. Inference

Following Abadie et al. (2010, 2015), we use placebo and permutation tests to examine whether the effect of holding the Olympics is statistically significant. Using these tests, we are able to evaluate whether the estimated effect for the treated unit is large relative to the effect estimated for a unit chosen at random. More specifically, first, we iteratively apply the SCM for each unit in the donor pool to estimate a placebo effect, assuming that the Olympics would be held in other countries at time $T_0 + 1$ (i.e., placebo-in-space tests). Second, as in Cho (2020) and Friedson et al. (2021), considering the quality of pre-treatment fit for the placebo effects, we calculate the country-specific RMSPE ratio of the post-treatment RMSPE to the pre-treatment RMSPE for each donor and then obtain a distribution of the RMSPE ratios for the placebo effects. Finally, we rank the RMSPE ratio for the actual treatment unit by comparing to the distribution of the RMSPE ratios and calculate a p-value for the significance of the treatment effect. When the RMSPE ratio in Tokyo ranks the first compared to the 37 donor countries, the two-sided p-value is 0.0263 (=1/38), indicating that the p-value cannot be less than 0.0263. Additionally, as in Mitze et al. (2020), we follow Altman and Bland (2011) to estimate the 90 percent confidence intervals from the two-sided p-values.

3.4. Ridge Augmented synthetic control method

As an application extension of the SCM, we use an Augmented synthetic control method (ASCM) recently developed by Ben-Michael et al. (2021) to estimate the impact of the Tokyo Olympics on COVID-19 cases. Abadie et al. (2010, 2015) suggest that if it is not possible to construct an SC that accurately fits the path of the actual outcome for the pre-treatment period, the estimated SC (counterfactual) should not be used for analysis. Ben-Michael et al. (2021) aim to correct the bias resulted from imbalance in pre-treatment outcomes between the treatment unit and the SC. The ASCM uses an outcome model to estimate and adjust the bias due to poor pre-treatment fit in the standard SCM estimate. Ben-Michael et al. (2021) propose the use of ridge regression as the outcome model. They call ASCM with a ridge outcome regression "Ridge ASCM" (RASCM).

This approach can improve pre-treatment fit by allowing for negative weights on some control units while minimizing the extrapolation from the convex hull by directly penalizing the distance from non-negative SCM weights. When using the RASCM, if there are control units with large negative weights, there may be risks of overfitting and bias resulted from extrapolation because Abadie et al. (2010, 2015) show that retaining the constraint on non-negative weights can prevent bias from extrapolating outside the convex hull of the control units.[16] Therefore, we check whether there are control units with large negative weights when using the RASCM. If there are, the estimated counterfactuals may not be reliable, even if the pre-treatment fit of the SC is better than the standard SCM.

3.5. Analysis procedure

As a baseline analysis, we first use the SCM, with the 30-day pre-treatment window from June 23 to July 22, and the level of the outcome to estimate the impact of holding the Tokyo Olympics on COVID-19 cases in Tokyo. We then conduct robustness checks using the full pre-treatment window from April 2 to July 22 or the log of the outcome. To confirm the robustness of the results using the SCM, we use

---

[13] As shown in Appendix A, although the SCs estimated from the RASCM had good pre-treatment fits compared to those from the SCM, there were many control units with large negative weights. Therefore, the RASCM estimates were not reliable.

[14] In applications of the SCM, the approach of using all pre-treatment outcomes has become popular recently because it improves the pre-treatment fit of the SC and mitigates concerns of choice of predictors such as p-hacking (Botosaru and Ferman, 2019). As shown by Kaul et al. (2022) and Botosaru and Ferman (2019), even if adding other covariates to all pre-treatment outcomes as predictors, optimization procedure by Abadie et al. (2010) used to estimate the SC will render all other covariates irrelevant because they are not used for predicting the outcome.

[15] For example, the pre-treatment window was 7 days for Friedson et al. (2021), 14 days for Mitze et al. (2020), 14 days for Breidenbach and Mitze (2022), 25 days for Cho (2020), and 28 days for Dave et al. (2021).

[16] As an empirical illustration of the RASCM, Ben-Michael et al. (2021) estimated the effect of the 2012 Kansas tax cut on log GSP per capita. From the weights of 49 control units used to construct the SC for Kansas when using all pre-treatment outcomes, there were no control units with negative weights greater than -0.1 (i.e., no large negative weights).

the RASCM for our analysis.[17] Additionally, we estimate the impact of the Olympics on COVID-19 cases in Japan as a whole by conducting the same analysis as above.

To evaluate the treatment effect (TE) of holding the Olympics on COVID-19 cases, we use the actual COVID-19 cases and the estimated counterfactuals from July 23 to August 22 (i.e., from the opening day of the Olympics to two weeks after the closing day) because there are approximately 7- to 14-day lags between infection and the reporting confirmed cases (Siordia, 2020; Chernozhukov et al., 2021).[18] Using the post-treatment period from July 23 to August 22, we calculate the average daily TE of holding the Olympics and the cumulative TE. When using the log of COVID-19 cases as the outcome, the estimates obtained from the SCMs are converted to the level of the outcomes (i.e., exp (SC)) to easily compare to the baseline results.

4. Results

4.1. SCM estimates for Tokyo

Fig. 2 shows the SCM estimates of the impact of holding the Tokyo Olympics on newly confirmed COVID-19 cases per one million people in Tokyo when using the level of the outcome.[19] Panel A reports the daily actual outcome and its SC. Panel B reports the difference between them, indicating the prediction errors before the treatment and the TE of the Olympics after the treatment. From panel A, we note that the estimated SC almost perfectly fits the development of the actual outcome, indicating that it can be used as a credible counterfactual. Panel B shows that the TE increased rapidly from a few days after the opening of the Olympics, and it became increasingly larger as time passed, reaching a maximum around August 19. The size of the effect increased to 154 cases on August 8 and 185 cases on August 22. Using the SCM estimates from July 23 to August 22, we calculate the average and cumulative TEs in Table 1. From column (1), the average daily TE was 113 (p-value=0.026), indicating that the Olympics significantly increased the number of COVID-19 cases by 113 cases per one million people per day on average. Column (2) shows that the Olympics significantly increased the cumulative number of COVID-19 cases by 3511 cases per one million people during the period. These results clearly show that holding the Olympics led to a rapid increase in COVID-19 infection in Tokyo.

To check the robustness of the baseline results, we redo our analysis using the full pre-treatment window. From panels C and D in Fig. 2, although the pre-treatment fit of the SC is worse than the baseline due to the much longer pre-treatment period, the SC has a still good fit. Similar to the baseline, the TE expanded rapidly from a few days after the opening of the Olympics. Columns (3) and (4) in Table 1 show that the average and cumulative TEs are slightly larger than the baseline levels and they are significant. Therefore, we note that the baseline results do not substantially change, even when using the full pre-treatment window.

We redo our analysis using the log of the outcome, and the results are shown in Fig. 3. In the case of using the 30-day pre-treatment window, the SC has a perfect pre-treatment fit in the period of the growing path of

COVID-19 cases just before the Olympics. Comparing columns (1)-(2) to (5)-(6) in Table 1, the average and cumulative TEs when using the log of the outcome are almost the same as those when using the level of the outcome. In the case of using the full pre-treatment window, panels C and D show that the SC has a poor pre-treatment fit compared to the SC estimated from using the level of the outcome, indicating that the SC in this case may not be reliable. However, columns (7) and (8) in Table 1 show that the average and cumulative TEs are not much different from those when using the level of outcome. Therefore, we note that our results do not substantially change, even when using the log of the outcome.

4.2. RASCM estimates for Tokyo

To confirm the robustness of the results using the SCM, we use the RASCM for our analysis in Figs. 4 and 5 and Table 2.[20] Compared to the SCs estimated from the SCM in the same analysis setting, the SCs estimated from the RASCM have good pre-treatment fits because the RASCM allows for negative weights on control units. As shown in Section 3.4, if there are control units with large negative weights even if the extrapolation from the convex hull is minimized by directly penalizing the distance from non-negative SCM weights, there may be risks of overfitting and bias resulted from extrapolation. Appendix C shows that for the RASCM when using the log of the outcome and the full pre-treatment window, there are 16 control units with negative weights, including 7 control units with negative weights greater than -0.1. Moreover, the SC in this case has the poorest pre-treatment fit in the RASCM analysis, even if large negative weights are placed on many control units. Therefore, the RASCM estimates in this case are not reliable (panels C and D in Fig. 5 and columns (7) and (8) in Table 2).

As shown in Figs. 4 and 5, the trajectories of the TEs estimated from the RASCM are similar to those estimated from the SCM: the TEs expanded rapidly from a few days after the start of the Olympics, and they reached their maximum around August 19. Table 2 shows that the average and cumulative TEs are slightly different and less significant compared to the SCM, but they are still significant and quite large, except for the less reliable result. The Olympics significantly increased the number of COVID-19 cases by 105 to 125 cases per one million people per day on average. Holding the Olympics led to a 3261 to 3871 increase in cumulative new COVID-19 cases per one million people during the period. Therefore, the RASCM results reinforce the robustness of the SCM results.

4.3. SCMs estimates for Japan

Additionally, we estimate the impact of the Tokyo Olympics on COVID-19 cases in Japan as a whole using the SCMs, although the main purpose of this paper is to estimate the impact of the Olympics on the host city of Tokyo. Tables 3 and 4 and Figs. 6–9 summarize the SCM and RASCM estimates for Japan. For the SCM results, Fig. 6 and 7 show that the SCs have good pre-treatment fits, except for the case of using the full pre-treatment window and the log of the outcome. As shown in these figures, as the Olympics progressed, the TEs in Japan increased rapidly. Table 3 shows that the average and cumulative TEs are significant and large, except for unreliable results due to a poor pre-treatment fit. For example, the Olympics significantly increased the number of COVID-19 cases in Japan by 47 to 65 cases per one million people per day on average.

Figs. 8 and 9 show that the SCs estimated from the RASCM have good pre-treatment fits compared to the SCs estimated from SCM in the same

---

[17] In addition to these tests, we performed several robustness checks in Section 5.

[18] Naturally, it is difficult to specify from when to when the Olympics had an influence. Considering the time lag between infection and the reporting confirmed cases, if persons were reported as COVID-19 patients on July 23, they were infected between July 9 and July 16. This means that our baseline estimates using the event day on July 23 may also include an infection of COVID-19 before the opening of the Olympics.

[19] In our implementation of the SCM, we used the Stata ado-file (*synth*) provided by Abadie et al. (2010).

[20] In our implementation of the RASCM, we used the R code (*augsynth*) provided by Ben-Michael et al. (2021).

analysis setting. Similar to the case of Tokyo, the RASCM estimates when using the log of the outcome and full pre-treatment window are not reliable because there are many control units with large negative weights.[21] As shown in Figs. 8 and 9, the trajectories of the TEs estimated from the RASCM are similar to those estimated from the SCM. Table 4 shows that the Olympics increased the number of COVID-19 cases in Japan by approximately 50 cases per one million people per day on average, except for unreliable or insignificant estimates.[22] Therefore, although the impact for Japan is much smaller than that for Tokyo, holding the Tokyo Olympics led to a large increase in the number of COVID-19 infections in Japan as a whole.

5. Additional robustness checks

To further increase the reliability of our results, we perform additional robustness checks in this section.[23] The detailed results are shown in the Appendix. First, we check whether our results are sensitive to excluding Australia from the donor pool because Australia had the maximum weight (more than 60%) in the SCMs when using the full pre-treatment window. Appendix D shows that our results do not substantially change, even when excluding Australia from the donor pool.

Second, we examine whether the baseline results change when using OECD countries with increasing spread of the delta variant of COVID-19 in July 2021 as the control units. In Japan, a more infectious delta variant spread in July, 2021. From biweekly data provided by the Our World in Data, the delta variant's share of SARS-CoV-2 sequences in Japan was 12.52% on June 28 and 61.07% on July 26, representing an increase of 48.55 percentage points in the month of July.[24] As an analysis that takes into account the spread of the delta variant in July, we conduct the SCMs using a donor pool of 17 OECD countries whose share of the delta variant increased more than Japan (i.e., an increase of more than 48.55 percentage points) in July.[25] Appendix E shows that the SCs have good pre-treatment fits and that the average and cumulative TEs are very similar to the baseline estimates. According to the context of the SCM, the results indicate that our method can construct reliable SCs (counterfactuals) controlled for the delta variant spread that affected COVID-19 cases in Tokyo and Japan, whether we use 37 or 17 countries as the control units.[26] However, it may not control for the delta variant if the prevalence of delta variant was more strongly affecting the actual COVID-19 cases only in Tokyo and Japan after the Olympics. If so, the estimated treatment effect of holding the Olympics may be overestimated. This is a limitation of our method.

Third, we examine how our estimates change when setting the event start date for the SCMs before July 23 to test whether there were substantial anticipation effects of holding the Olympics (i.e., placebo-in-time tests). Specifically, we redo our analysis by initially setting the event start date to July 9 (i.e., the day after announcing that the Olympics would be held without spectators) and then changing the event start date from July 9 to every other day.[27] Appendix F shows that there were small but no sizable deviations of the actual outcomes from the SCs before July 23, indicating that there were no substantial anticipation effects.[28]

Fourth, given the time lags between infection and the reporting confirmed cases, we examine how the SCMs estimates change if the evaluation period is changed from July 23-August 22 to July 30-August 22. From Appendix G, the average TEs in Tokyo (Japan) were approximately 30 cases (approximately 15 cases) larger than the baseline estimates. Due to the shorter evaluation period, the cumulative TEs were slightly smaller than the baselines, but not much different from them. Therefore, the results support the evidence that holding the Olympics led to a rapid increase in COVID-19 infection in Tokyo (Japan).

Fifth, taking the stationarity (non-stationarity) of the data into account, we examine how the baseline estimates change when using the daily change rate of COVID-19 cases as the outcome to be matched in the SCMs. To compare to the baseline estimates, we first set that on July 22 (i.e., the last day before the opening of the Olympics), the level of the SC is equal to the actual level of COVID-19 cases. Then, using July 22 as a benchmark point, we calculate the level of the counterfactual by piecing together the estimated SC of the daily change rate from July 22 to August 22. Finally, we estimate the daily average and cumulative TEs of holding the Olympics. Appendix H shows that the trajectories of the TEs and the average TEs are similar to the baselines.[29] Therefore, our baseline estimates are robust to using the daily change rate of COVID-19 cases as the outcome.

Sixth and finally, we estimate the impact of holding the Olympics on effective reproduction number of COVID-19 using the SCMs. The effective reproduction number (R_t) is used as a leading indicator of the extent of new infection spread. If R_t is greater than 1, it indicates that the number of new cases will increase. Following the methodology of the Toyo Keizai Online, we simply calculate an R_t for Tokyo and 38 OECD countries.[30] From Appendix I, the SCs have good pre-treatment fits, indicating that they can be used as reliable counterfactuals.[31] Both of the actual R_t and the SC were greater than 1 in the evaluation period. The estimated TEs (i.e., deviations of the actual R_t from the SCs) increased from July 27, reached a maximum around August 1, and decreased from

---

[21] In this case, there were 16 control units with negative weights, including 7 control units with negative weights greater than -0.1 (see Appendix C).

[22] The SCM estimates for Japan using the level of the outcome and the full pre-treatment window were significant, while the RASCM estimates using the same setting were not significant.

[23] Following the referees' comments, we performed some tests presented in this section. We are grateful to two anonymous referees because these tests helped to increase the validity of our results.

[24] The data were obtained on March 12, 2022.

[25] The 17 countries were as follows: Belgium, Canada, Czechia, Denmark, France, Ireland, Italy, Latvia, Lithuania, Mexico, Netherlands, Norway, Poland, Slovakia, Slovenia, Sweden, and Switzerland.

[26] In the context of the SCM, if we can construct an SC that correctly fits the actual outcome in the pre-treatment period, we can control for time-varying observed and unobserved factors (e.g., delta variant) that affect the outcome and the treatment.

[27] For this test, the full pre-treatment window was used to ensure the length of the pre-sample for the SCMs.

[28] If we consider small deviations of the actual outcomes from the SCs as anticipation effects, following the advice of Abadie et al. (2010, 2015), we need to set the event start date before July 23. In that case, the counterfactuals were slightly smaller than the baselines because the SCs could not be constructed to match the expansion of COVID-19 cases just before the Olympics.

[29] In the case of using the full pre-treatment window, the average TE in Tokyo (Japan) was 114 (53) for the SCM and 106 (48) for the RASCM, while the detailed results were omitted for brevity.

[30] According to the Toyo Keizai Online (https://toyokeizai.net/sp/visual/tko/covid19/), R_t = (Number of new COVID-19 cases in past 7 days / Number of new COVID-19 cases in 7 days before that) ^ (Mean generation time / Length of reporting interval), where mean generation time is 5 and length of reporting interval is 7.

[31] Although we estimated the impact of the Olympics on R_t using the full pre-treatment window, we could not obtain reliable counterfactuals because the SCs had poor pre-treatment fits.

around August 4. The average TE in Tokyo (Japan) was 0.18 (0.20) for the SCMs.[32] Therefore, these results in Appendix I show that holding the Olympics spread COVID-19 infection in Tokyo and Japan.

6. Discussion and conclusion

We used the SCMs to estimate the impact of holding the Tokyo Olympics on the number of newly confirmed COVID-19 cases in Tokyo (Japan). Using reliable estimates obtained from the SCMs,[33] we found that holding the Olympics significantly increased the daily average of new COVID-19 cases by 105 to 132 cases in Tokyo (47 to 65 cases in Japan) per one million people from July 23 to August 22 compared to the counterfactuals. From these estimates and the population of Tokyo and Japan, we calculated that if the Olympics had not been held, the average daily number of COVID-19 cases could have been reduced by as many as approximately 1500 to 1850 cases in Tokyo and approximately 5900 to 8150 cases in Japan as a whole during that time period. These impacts were quite sizable. We may therefore reasonably conclude that holding the Tokyo Olympics was likely a factor in the spread of COVID-19 infection in the host city of Tokyo.

Our study has some limitations. Our research approach did not enable us to determine the mechanisms behind the results obtained. We offer a few conjectures. It is possible that the coronavirus could have been transmitted directly from the participants of the Olympics to the people of Tokyo. However, since there were no spectators in the venues around Tokyo and if the TOCOG implemented effective measures to combat COVID-19 infection, we believe that the number of new COVID-19 cases in Tokyo would not have increased as much as it did through this route. It may be hard to argue that the Olympics increased people's mobility and consequently the rapid spread of COVID-19 in Tokyo because the fourth state of emergency did not change people's mobility, and the mobility during and after the Olympics did not change substantially compared to last year, although it declined (see Appendix J).[34] We conjecture that holding the Olympics, which was an inconsistent policy with the state of emergency for preventing COVID-19, may have indirectly caused a decline in the motivation for self-restraint and awareness of prevention, resulting in a decline in infectious disease control measures at the individual level (Fujii et al., 2021). For example, according to a poll conducted by Asahi Shimbun (2021b) on August 7 and 8, 61% of respondents said that holding the Tokyo Olympics "had loosened" the mood of self-restraint in society against COVID-19 infection.

Our study is limited by the quality of the data.[35] It is based on observational data, rather than data from laboratory experiments or from social experiments with randomized controlled trials (RCTs). It cannot control for all external conditions that may influence the impact of the Olympics on COVID-19 cases.[36] However, the SCM is a valuable approach for evaluating the causal impact on population-level health outcomes when an RCT is impractical, such as the Olympics (e.g., Bouttell et al., 2018). Using the SCMs, we succeeded in estimating reliable counterfactuals controlled for time-varying observed and unobserved factors that affected the outcome and the treatment by constructing the SCs that accurately approximated the trajectories of COVID-19 cases for Tokyo or Japan in the pre-treatment period. However, our method using the SCMs cannot fully control for all other confounding factors specific only to Tokyo or Japan after the opening of the Olympics. Hence, the estimated treatment effect of holding the Olympics could be biased if there were specific factors that strongly influenced the actual COVID-19 cases only in Tokyo or Japan after the Olympics. This is a limitation of our method.

Since our research is a case study of the impact of holding the Tokyo Olympics on COVID-19 cases in the host city of Tokyo during the COVID-19 pandemic, it may not be possible to generalize our empirical results from this study. Holding the Olympics itself is a rare event, and moreover, the Olympics have never before been held in the midst of a new virus spread. Although we conducted a case study of the ultimately rare event, we believe that our findings provide useful information for understanding the relationship between large-scale national events such as the Olympics and infectious diseases.

Conflict of Interest

We declare no conflict of interest.

Acknowledgements

Appendix A. SCMs results for Tokyo when using Japanese prefectures as the control units

Figs. A1, A2, A3

---

[32] In Appendix I, although the cumulative TEs were shown in a way that fitted the other figures, they might not be appropriate in evaluating the effective reproduction number.

[33] As shown in Section 4, the estimates obtained from using the log of the outcome and the full pre-treatment window were not reliable.

[34] Watanabe and Yabu (2021a, 2021b) show that people in Japan responded to an increase in new COVID-19 cases in 2020 by voluntarily refraining from going out. Hosono, 2021 shows that such a voluntary lockdown accounts for a major part of the low COVID-19 infection rates in 2020 in Japan. While Watanabe and Yabu (2021a, 2021b) and Hosono, 2021 stress the demand-side effects of voluntary restriction on movement, Hayakawa et al. (2022) discuss the supply-side effects of restaurant restraining order in Japan. If the voluntary lockdown had occurred after the opening of the Olympics, and if it had been effective in preventing the spread of COVID-19 infection, the number of new COVID-19 cases in Tokyo (Japan) after the Olympics might have been much larger than there would have been without the voluntary lockdown. In addition, if the degree of the voluntary lockdown was higher in Japan than in other OECD countries, the SC for Tokyo (Japan) constructed as a weighted average of COVID-19 cases of the 37 OECD countries may be overestimated. Thus, if the above two cases hold, the estimated treatment effect of holding the Olympics may be underestimated.

[35] As is common in research on COVID-19, the reported number of newly confirmed COVID-19 cases is understated compared to the true cases because not all people have been tested.

[36] As in conventional causal inference methods, if there are other events after the treatment that would affect the outcome only in the treatment unit, the SCM cannot correctly identify the effects of treatment and such events. For example, since the fourth state of emergency had been declared in Tokyo since July 12, the Tokyo Olympics were held from July 23 to August 8 under the state of emergency. The analysis by setting the event start date for the SCM to July 9 and 11 shows that the infection control effect of the state of emergency was not seen before the Olympics (see Appendix F). However, if the infection control effect of the state of emergency had been substantial after the opening of the Olympics, the number of new COVID-19 cases in Tokyo after the Olympics might have been much larger than in the absence of the state of emergency. Thus, if such is the case, the estimated treatment effect of holding the Olympics may be underestimated.
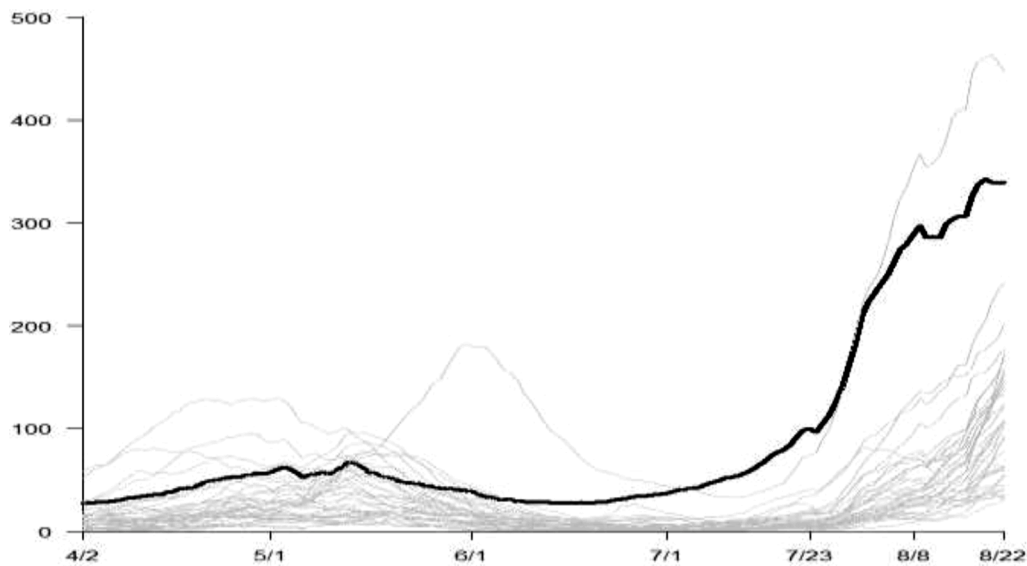
**Fig. A1.** COVID-19 cases in Japan by prefecture.

Notes: Fig. A1 shows the development of daily 7-day moving average of newly confirmed COVID-19 cases per million people by each Japanese prefecture. The control units include 37 Japanese prefectures, excluding Tokyo, Kanagawa, Saitama, Chiba, Ibaraki, Yamanashi, Shizuoka, Miyagi, Fukushima, and Hokkaido (which have Olympic venues or are in close proximity to Tokyo). The number of COVID-19 cases for Tokyo is shown in the thick line and those for other prefectures are shown in the thin lines.

Sources: Japanese Ministry of Health, Labour and Welfare and Our World in Data.
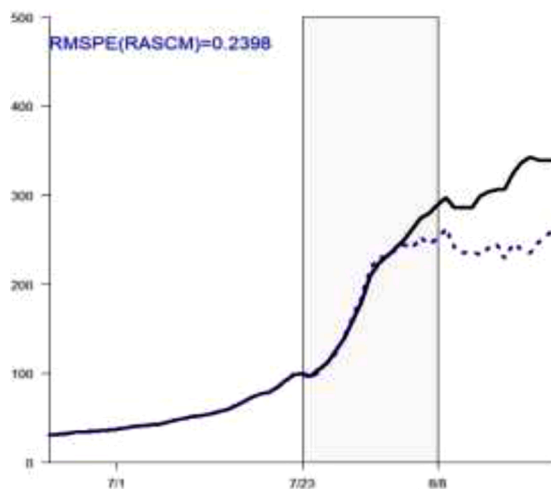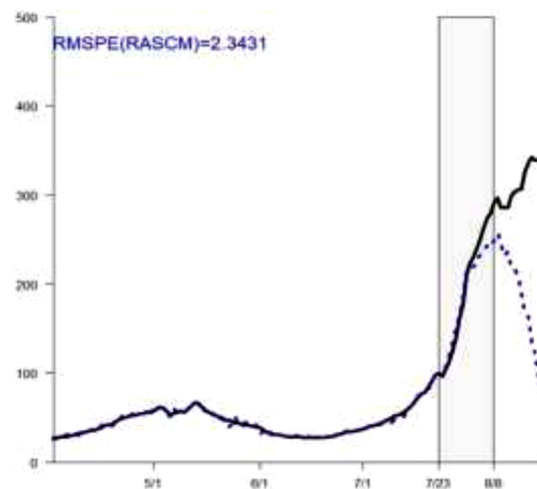
A. Actual vs. SC: SCM, pre-window=30 days
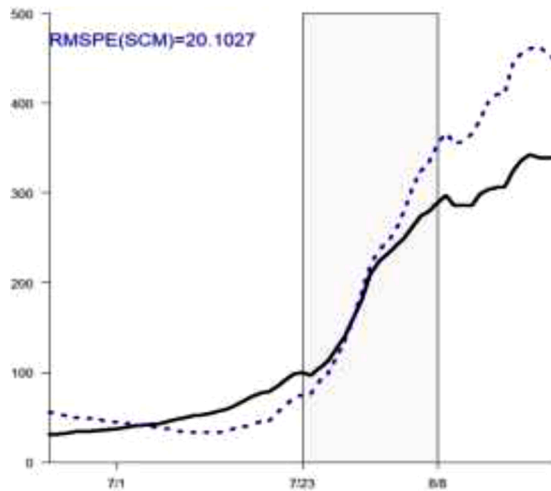
B. Actual vs. SC: SCM, pre-window=full
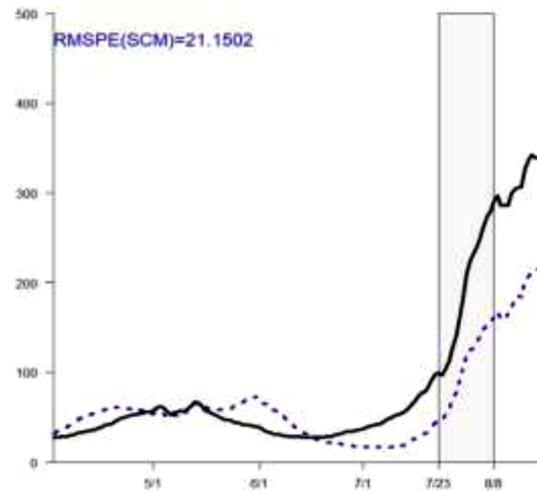

C. Actual vs. SC: RASCM, pre-window=30 days

D. Actual vs. SC: RASCM, pre-window=full

Fig. A2. SCMs estimates for Tokyo: using Japanese prefectures as the control units.

Notes: Fig. A2 shows the SCMs estimates for Tokyo when using the level of the outcome and Japanese prefectures as control units. The control units include 37 Japanese prefectures, excluding Tokyo, Kanagawa, Saitama, Chiba, Ibaraki, Yamanashi, Shizuoka, Miyagi, Fukushima, and Hokkaido (which have Olympic venues or are in close proximity to Tokyo). Panels A and B show the SCM estimates for Tokyo. Panels C and D show the RASCM estimates for Tokyo. The development of the actual COVID-19 cases is shown in the black solid line and that of the SCs is shown in the blue dashed line. The shadow areas represent the duration of the Olympic.

A. Outcome=level, pre-window=30 days   B. Outcome=level, pre-window=full
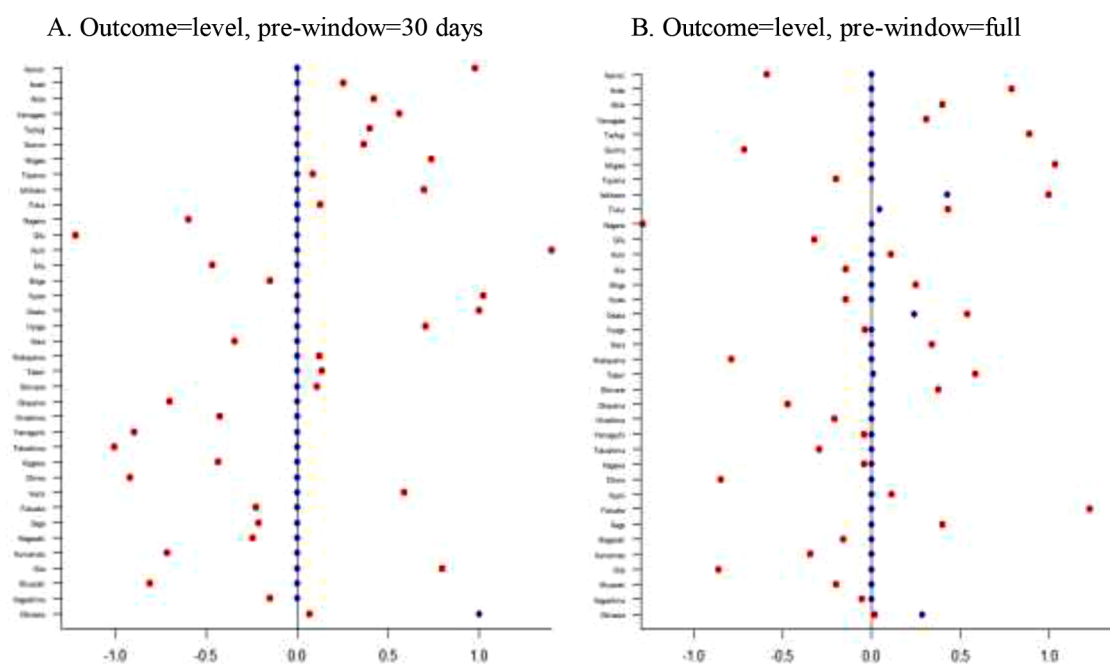
Fig. A3. Weights for control units used to construct the SC for Tokyo.
Notes: Fig. A3 shows the estimated weights for 37 Japanese prefectures used to construct the SC for Tokyo by using the SCMs. The weights estimated from the SCM are shown in the blue dots, and the weights estimated from the RASCM are shown in the red dots.

Appendix B. Comparison COVID-19 cases in Tokyo (Japan) to those in 37 OECD countries
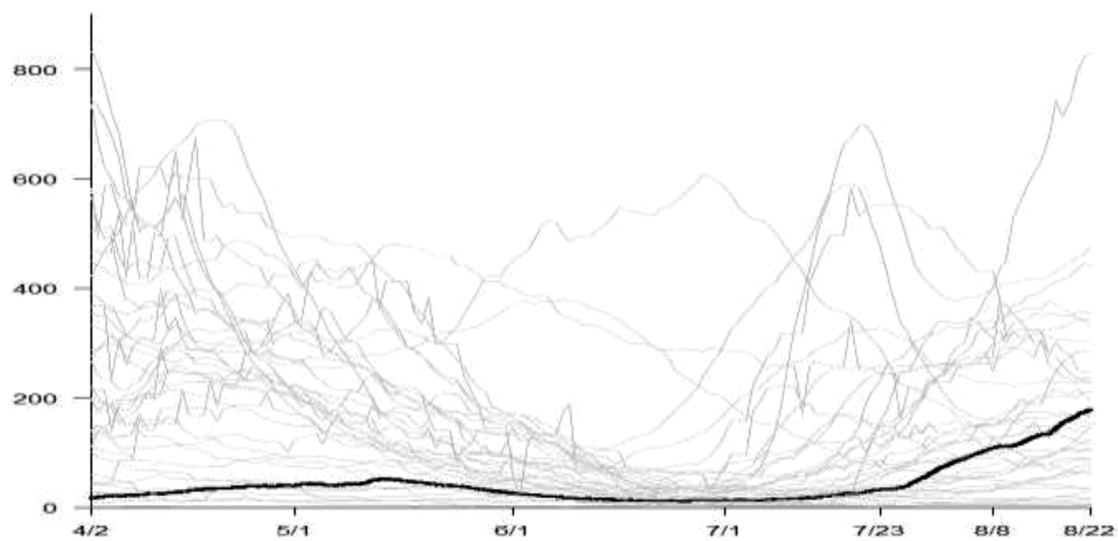
Figs. B1, B2
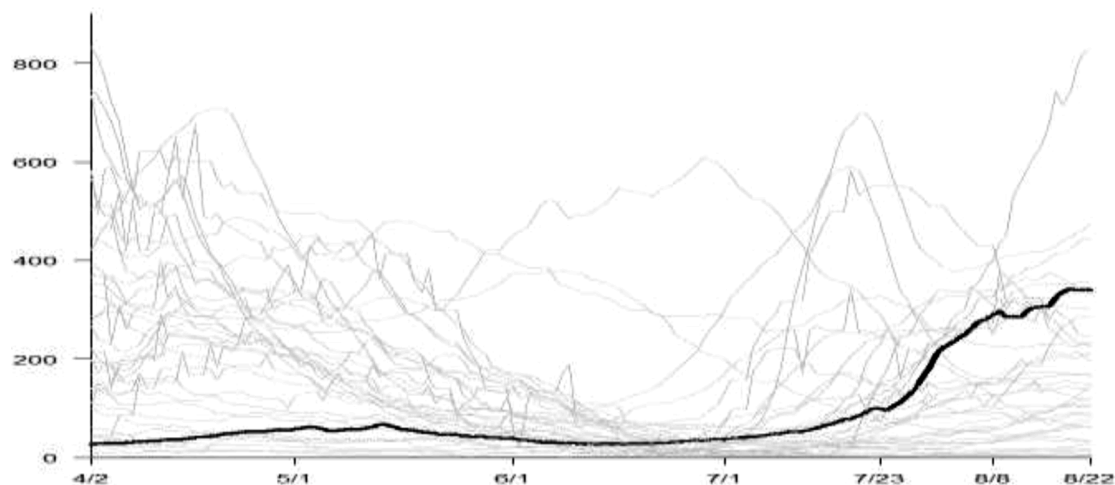
Fig. B1. COVID-19 cases between Tokyo and 37 OECD countries.
Notes: Fig. B1 shows the development of daily 7-day moving average of newly confirmed COVID-19 cases per million people between Tokyo and 37 OECD countries. The number of COVID-19 cases for Tokyo is shown in the thick line and those for 37 OECD countries are shown in the thin lines.
Sources: Japanese Ministry of Health, Labour and Welfare and Our World in Data.
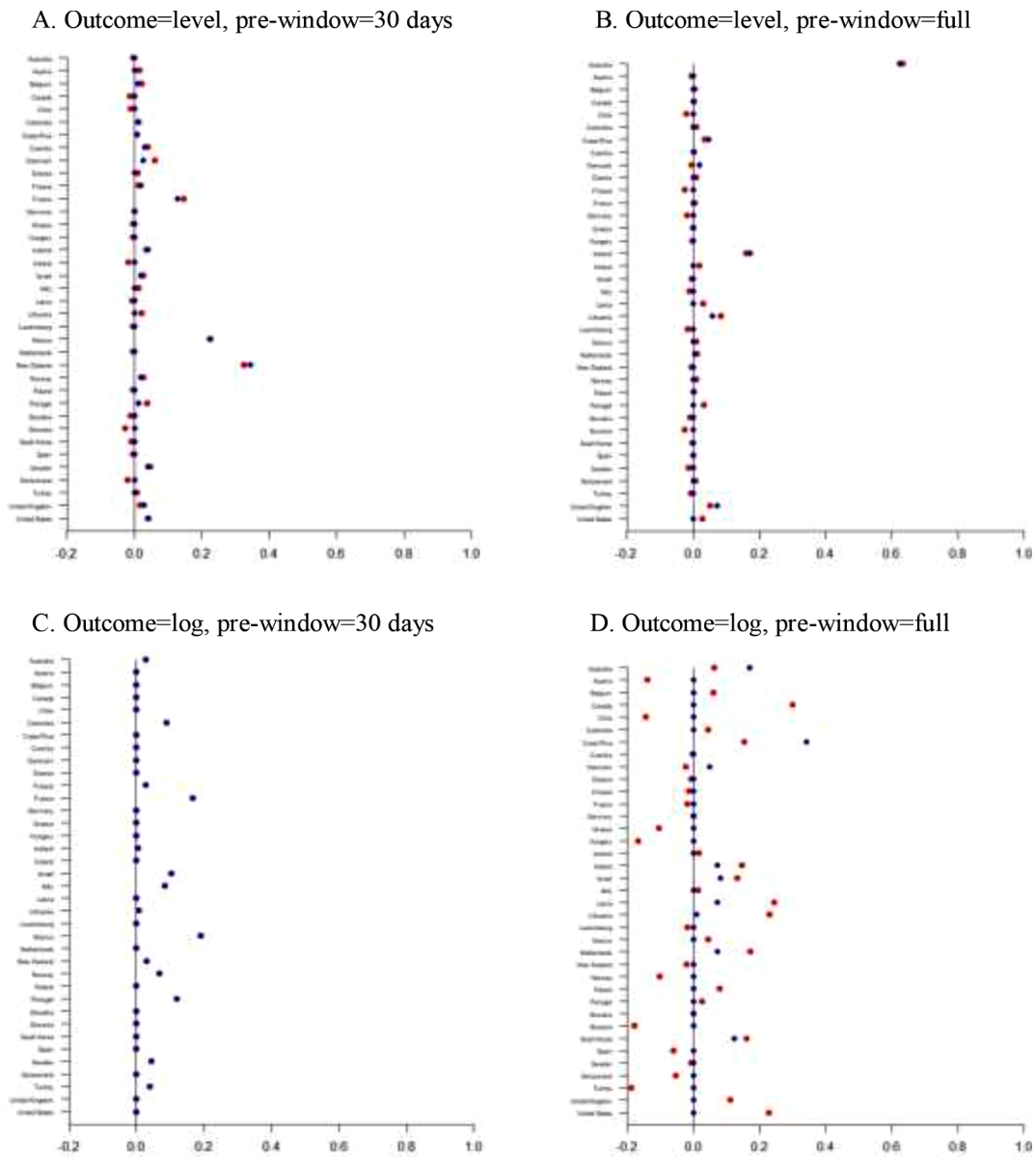
Fig. B2. COVID-19 cases between Japan and 37 OECD countries.
Notes: Fig. B2 shows the development of daily 7-day moving average of newly confirmed COVID-19 cases per million people between Japan and 37 OECD countries.
The number of COVID-19 cases for Japan is shown in the thick line and those for 37 OECD countries are shown in the thin lines.
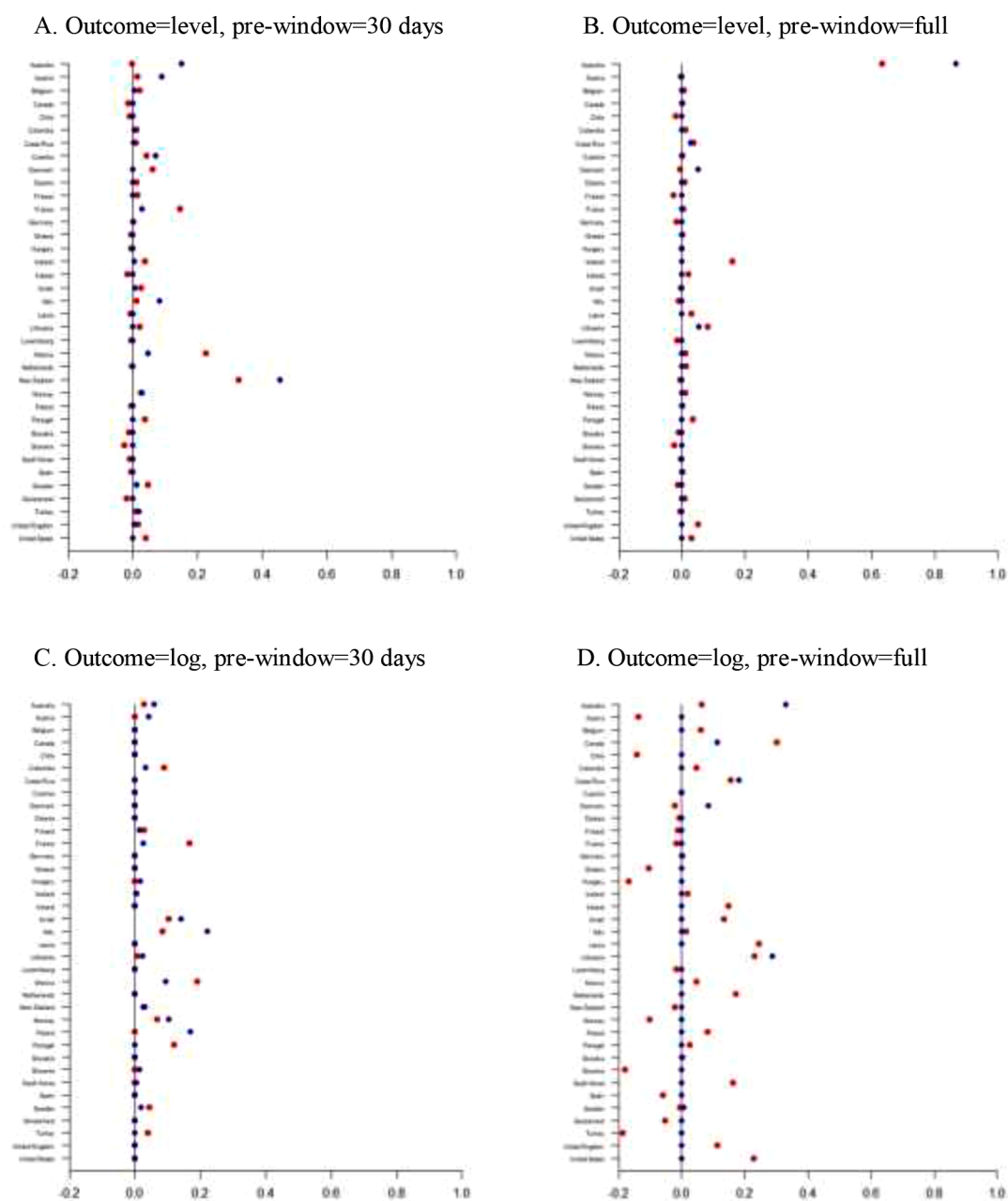Sources: Japanese Ministry of Health, Labour and Welfare and Our World in Data.

Appendix C. Weights for the control units of 37 countries

Figs. C1, C2

A. Outcome=level, pre-window=30 days

B. Outcome=level, pre-window=full

C. Outcome=log, pre-window=30 days

D. Outcome=log, pre-window=full

Fig. C1. Weights for control units used to construct the SC for Tokyo.
Notes: Fig. C1 shows the estimated weights for 37 OECD countries used to construct the SC for Tokyo by using the SCMs. The weights estimated from the SCM are shown in the blue dots, and the weights estimated from the RASCM are shown in the red dots.

A. Outcome=level, pre-window=30 days

B. Outcome=level, pre-window=full

C. Outcome=log, pre-window=30 days

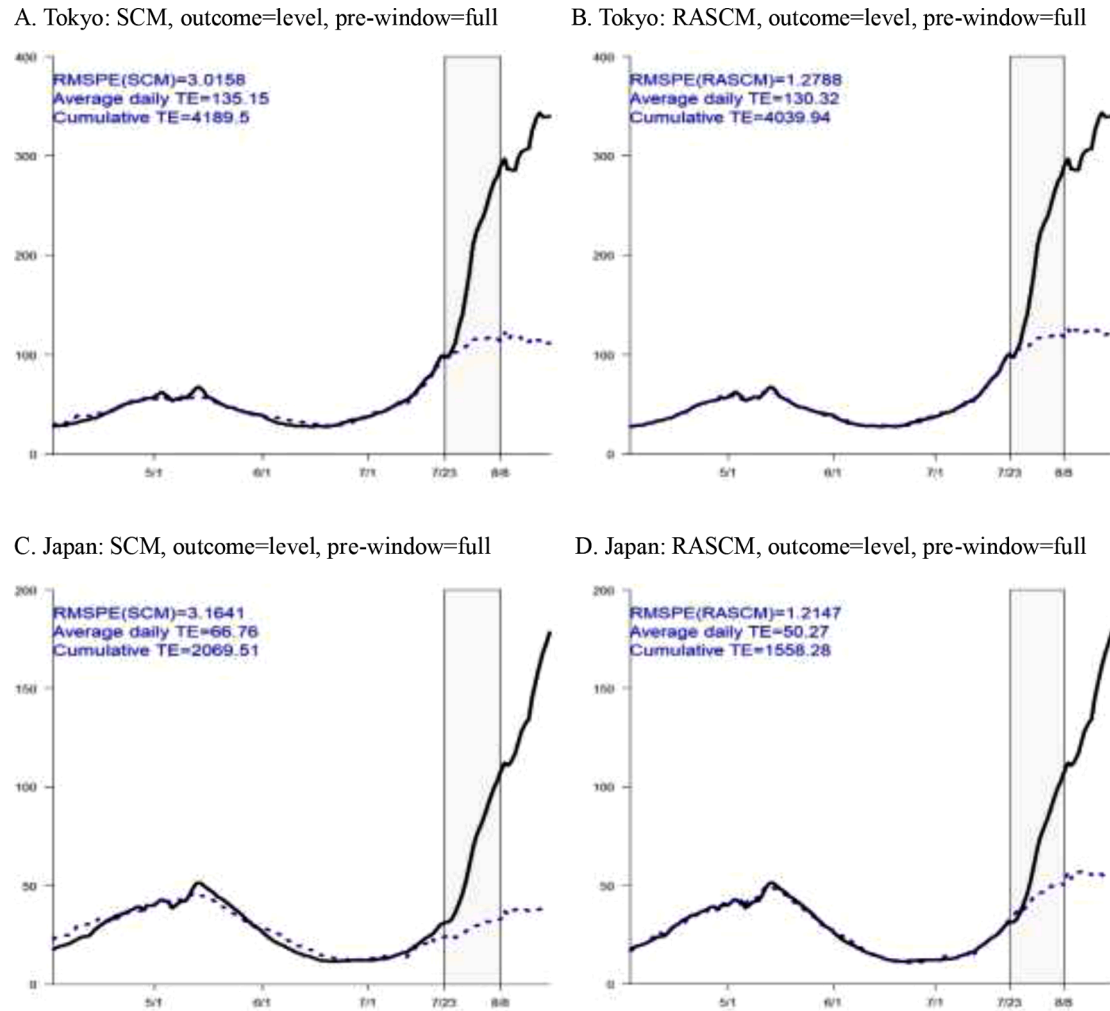D. Outcome=log, pre-window=full

Fig. C2. Weights for control units used to construct the SC for Japan.
Notes: Fig. C2 shows the estimated weights for 37 OECD countries used to construct the SC for Japan by using the SCMs. The weights estimated from the SCM are shown in the blue dots, and the weights estimated from the RASCM are shown in the red dots.
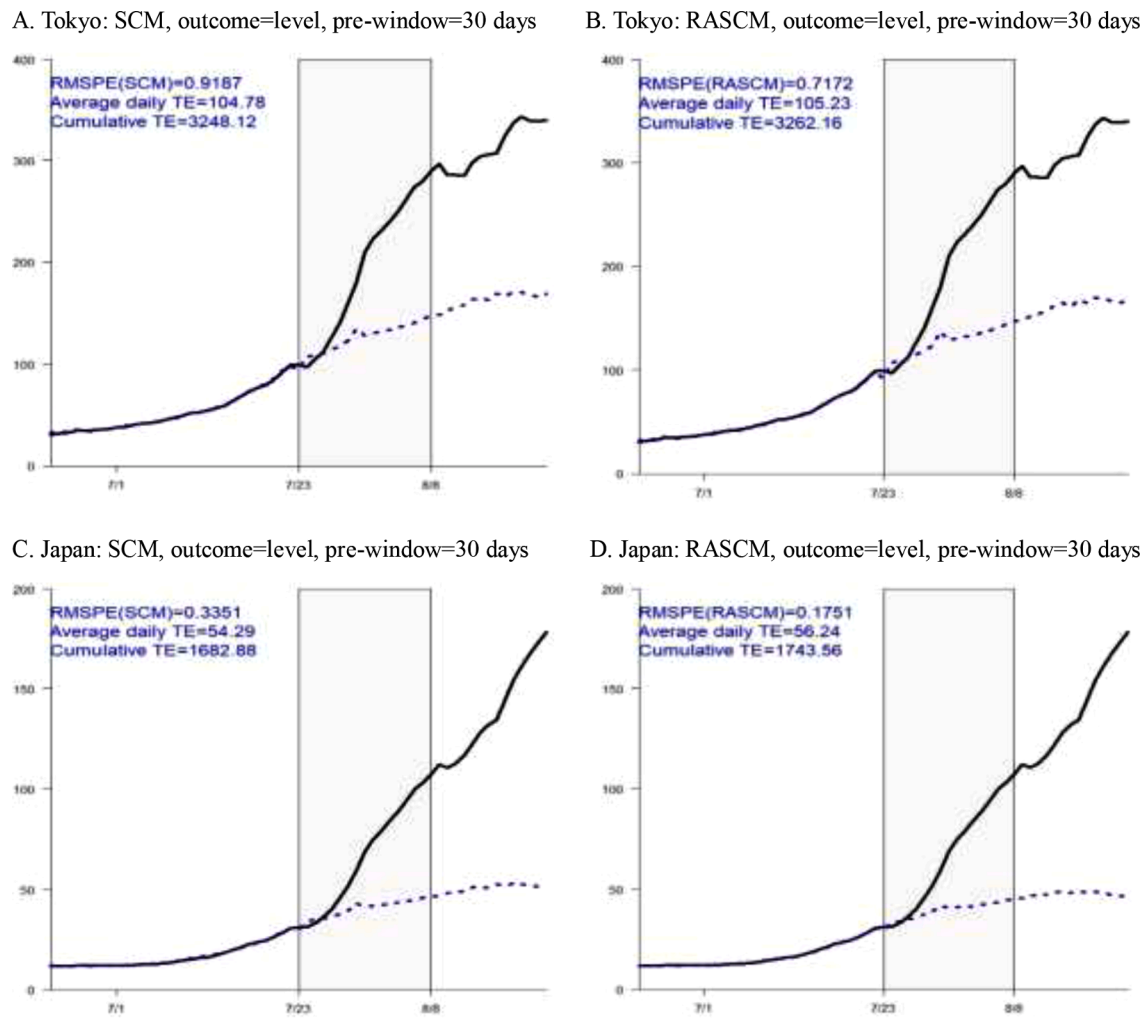
Appendix D. Excluding Australia from the donor pool

Fig. D

A. Tokyo: SCM, outcome=level, pre-window=full

RMSPE(SCM)=3.0158
Average daily TE=135.15
Cumulative TE=4189.5

B. Tokyo: RASCM, outcome=level, pre-window=full

RMSPE(RASCM)=1.2788
Average daily TE=130.32
Cumulative TE=4039.94

C. Japan: SCM, outcome=level, pre-window=full

RMSPE(SCM)=3.1641
Average daily TE=66.76
Cumulative TE=2069.51

D. Japan: RASCM, outcome=level, pre-window=full

RMSPE(RASCM)=1.2147
Average daily TE=50.27
Cumulative TE=1558.28

Fig. D. SCMs estimates for Tokyo and Japan: excluding Australia from the donor pool.
Notes: Fig. D shows the actual level of COVID-19 cases and the counterfactuals in Tokyo and Japan when excluding Australia (which had the maximum weight in the SCM estimates when using the full pre-treatment window) from the donor pool. The SCM and RASCM estimates for Tokyo are shown in panels A and B, and those for Japan are shown in panels C and D.

Appendix E. Using only countries with spread of delta variant of COVID-19 in July 2021 as the control units

Fig. E

A. Tokyo: SCM, outcome=level, pre-window=30 days

B. Tokyo: RASCM, outcome=level, pre-window=30 days

C. Japan: SCM, outcome=level, pre-window=30 days

D. Japan: RASCM, outcome=level, pre-window=30 days

Fig. E. SCMs estimates for Tokyo and Japan: using 17 OECD countries with spread of delta variant in July 2021 as the control units.

Notes: Fig. E shows the actual level of COVID-19 cases and the counterfactuals in Tokyo and Japan when using 17 OECD countries with spread of delta variant of COVID-19 in July 2021 as the control units. The 17 countries are as follows: Belgium, Canada, Czechia, Denmark, France, Ireland, Italy, Latvia, Lithuania, Mexico, Netherlands, Norway, Poland, Slovakia, Slovenia, Sweden, and Switzerland. The SCM and RASCM estimates for Tokyo are shown in panels A and B, and those for Japan are shown in panels C and D.
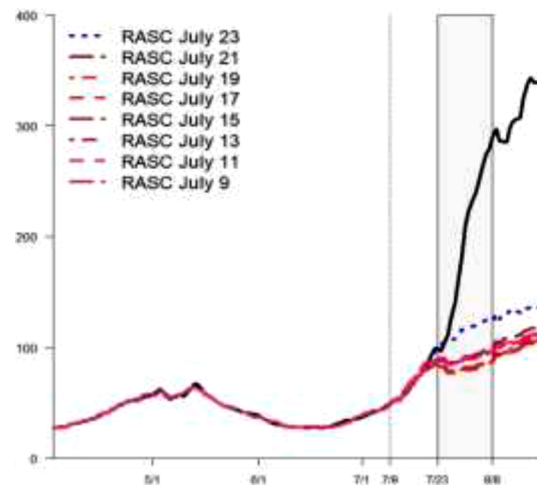
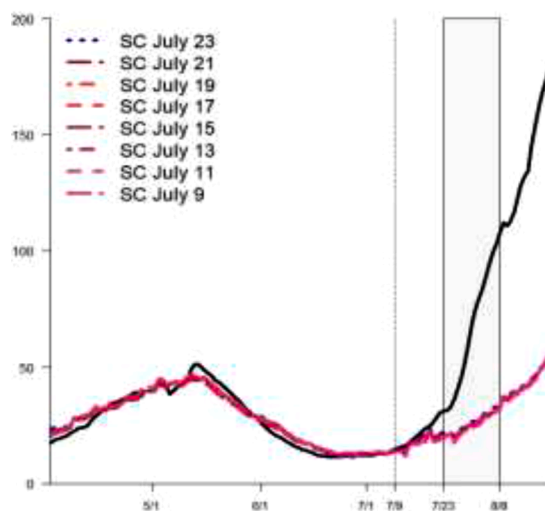**Appendix F. Placebo-in-time tests**

Fig. F

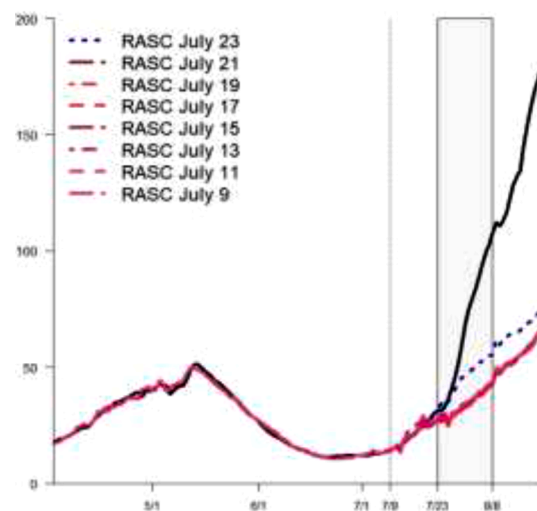A. Tokyo: SCM, outcome=level, pre-window=full
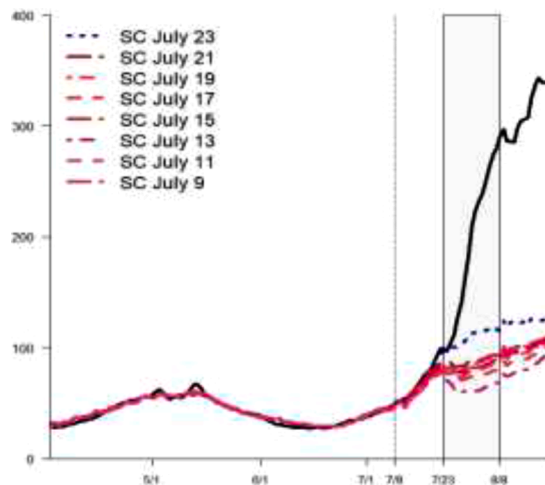
B. Tokyo: RASCM, outcome=level, pre-window=full

C. Japan: SCM, outcome=level, pre-window=full

D. Japan: RASCM, outcome=level, pre-window=full

Fig. F. Placebo-in-time tests: Tokyo and Japan.
Notes: We examine how the baseline results change when setting the event start date in the SCMs before July 23 (i.e., placebo-in-time tests). Specifically, we redo our analysis by changing the event start date from July 9 to every other day. The full pre-treatment window is used to ensure the length of the pre-sample for the SCMs. The SCM and RASCM estimates for Tokyo are shown in panels A and B, and those for Japan are shown in panels C and D.

Appendix G. Changing the evaluation period from July 23-August 22 to July 30-August 22

Table G1, G2

Table G1
SCMs estimates for Tokyo: changing evaluation period.

RMSPE=0.5649	(1)	(2)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	113.25	3510.78
July 30-August 22	143.22	3437.30
RMSPE=2.7851	(3)	(4)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	131.87	4087.88
July 30-August 22	164.73	3953.40
RMSPE=0.2319	(5)	(6)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	105.18	3260.69
July 30-August 22	133.84	3212.28
RMSPE=1.2760	(7)	(8)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	124.86	3870.66
July 30-August 22	157.11	3770.72

Notes: We examine how the baseline results change when the evaluation period is changed from July 23-August 22 to July 30-August 22. From the SCM and RASCM results, we calculate the average daily treatment effect (TE) and the cumulative TE in Tokyo.

Table G2
SCMs estimates for Japan: changing evaluation period.

RMSPE=0.3495	(1)	(2)
Method	SCM	SCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	52.75	1635.22
July 30-August 22	67.12	1610.86
RMSPE= 3.0003	(3)	(4)
Method	SCM	SCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	64.67	2004.71
July 30-August 22	78.83	1891.80
RMSPE=0.0848	(5)	(6)
Method	RASCM	RASCM
Pre-window	June 23 to July 22	June 23 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	50.83	1575.64
July 30-August 22	64.47	1547.26
RMSPE=1.1893	(7)	(8)
Method	RASCM	RASCM
Pre-window	April 2 to July 22	April 2 to July 22
Outcome	Level (cases)	Level (cases)
Evaluation periods and TE	Average daily TE	Cumulative TE
July 23-August 22	44.55	1380.92
July 30-August 22	57.12	1370.79

Notes: We examine how the baseline results change when the evaluation period is changed from July 23-August 22 to July 30-August 22. From the SCM and RASCM results, we calculate the average daily treatment effect (TE) and the cumulative TE in Japan.
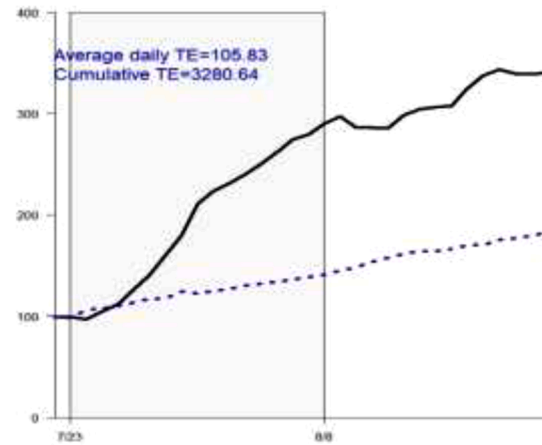
Appendix H. Using daily change rate of COVID-19 cases as the outcome
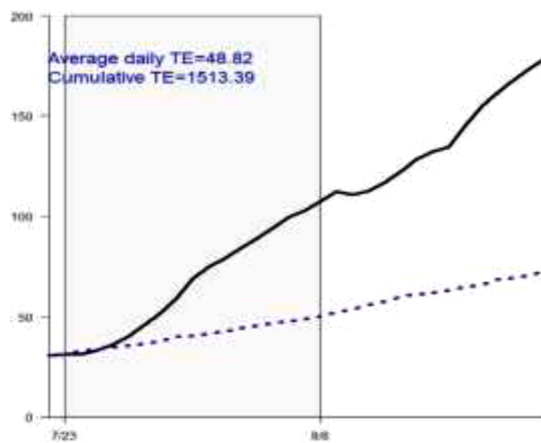
Fig. H

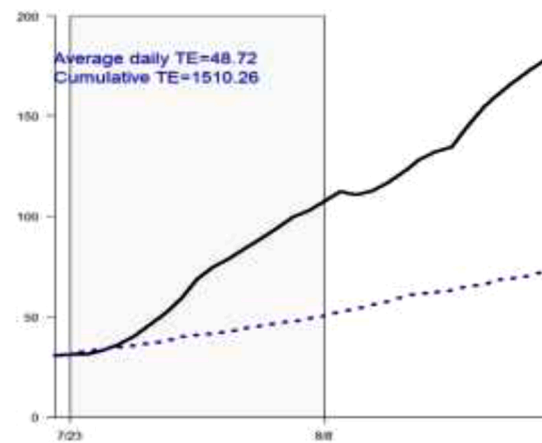A. Tokyo: SCM, outcome=level, pre-window=30 days   B. Tokyo: RASCM, outcome=level, pre-window=30 days
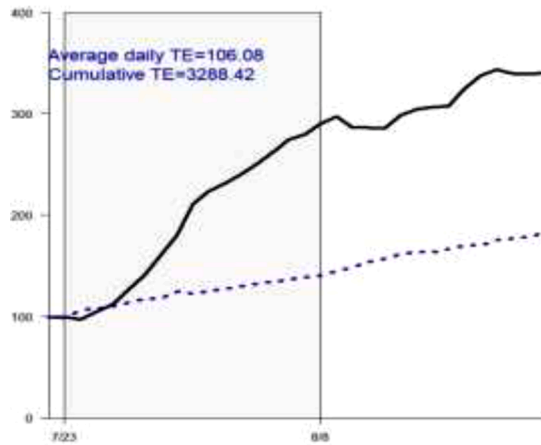
C. Japan: SCM, outcome=level, pre-window=30 days   D. Japan: RASCM, outcome=level, pre-window=30 days
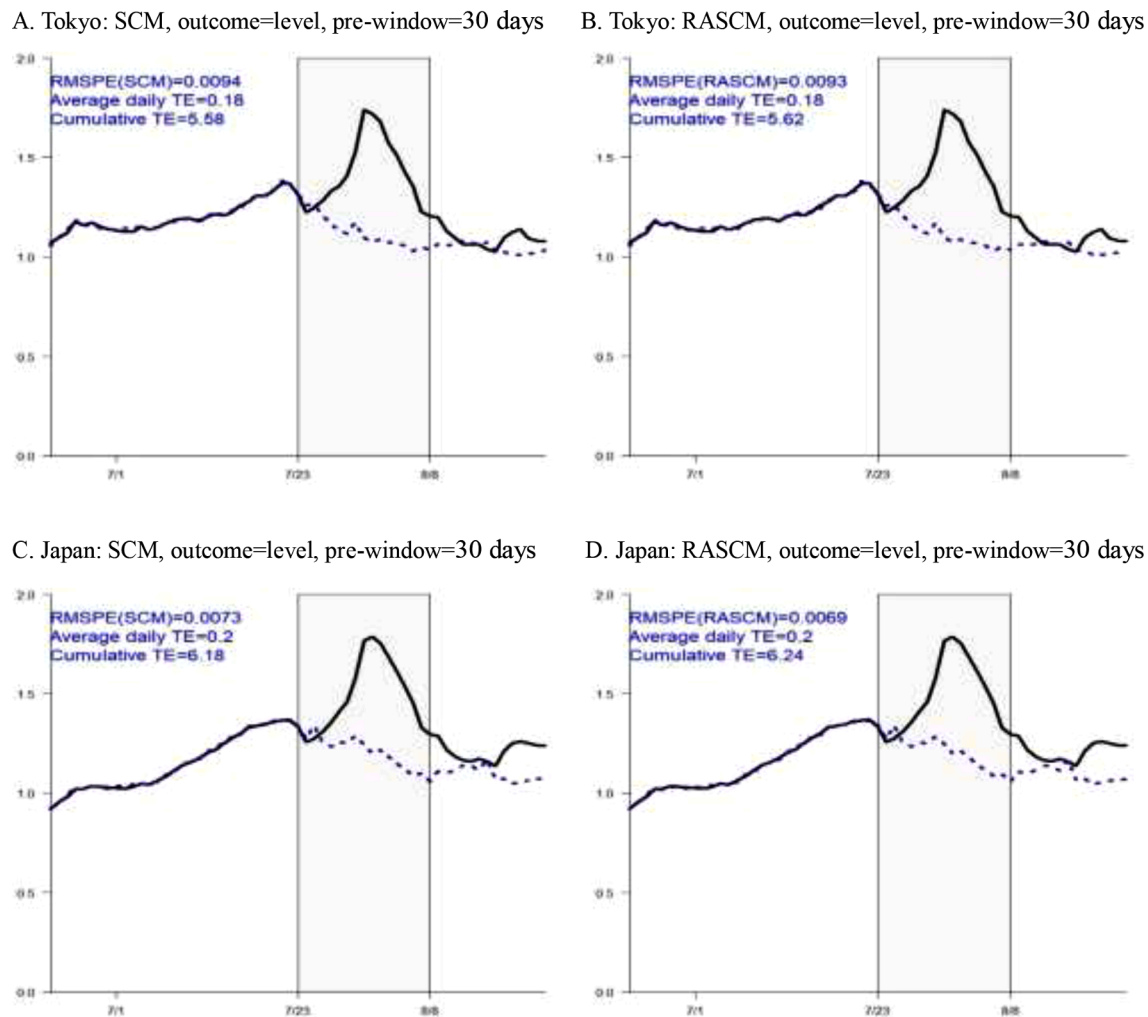

Fig. H. SCM estimates for Tokyo and Japan: daily change rate of COVID-19 cases.
Notes: In Fig. 1, taking the stationarity (non-stationarity) of the data into account, we present the actual level of COVID-19 cases and the counterfactual in Tokyo and Japan from July 22 to August 22 when using the daily change rate of COVID-19 cases as the outcome. To compare to the baseline estimates, using July 22 as a benchmark point, we calculate the level of the counterfactual by piecing together the estimated SC of the daily change rate from July 22 to August 22. The SCM and RASCM estimates for Tokyo are shown in panels A and B, and those for Japan are shown in panels C and D.

Appendix I. Using the effective reproduction number of COVID-19 as the outcome

Fig. I

A. Tokyo: SCM, outcome=level, pre-window=30 days

B. Tokyo: RASCM, outcome=level, pre-window=30 days

C. Japan: SCM, outcome=level, pre-window=30 days

D. Japan: RASCM, outcome=level, pre-window=30 days

Fig. I. SCMs estimates for Tokyo and Japan: effective reproduction number.

Notes: Fig. I shows the actual effective reproduction number of COVID-19 and the counterfactuals in Tokyo and Japan. Following the methodology of the Toyo Keizai Online (https://toyokeizai.net/sp/visual/tko/covid19/), we calculated an effective reproduction number (R_t) as follows: $R_t = \left(\frac{\text{Number of new COVID-19 cases in past 7 days}}{\text{Number of new COVID-19 cases in 7 days before that}} \right)^{\left(\frac{\text{Mean generation time}}{\text{Length of reporting interval}} \right)}$, where mean generation time is 5 and length of reporting interval is 7. The SCM and RASCM estimates for Tokyo are shown in panels A and B, and those for Japan are shown in panels C and D.
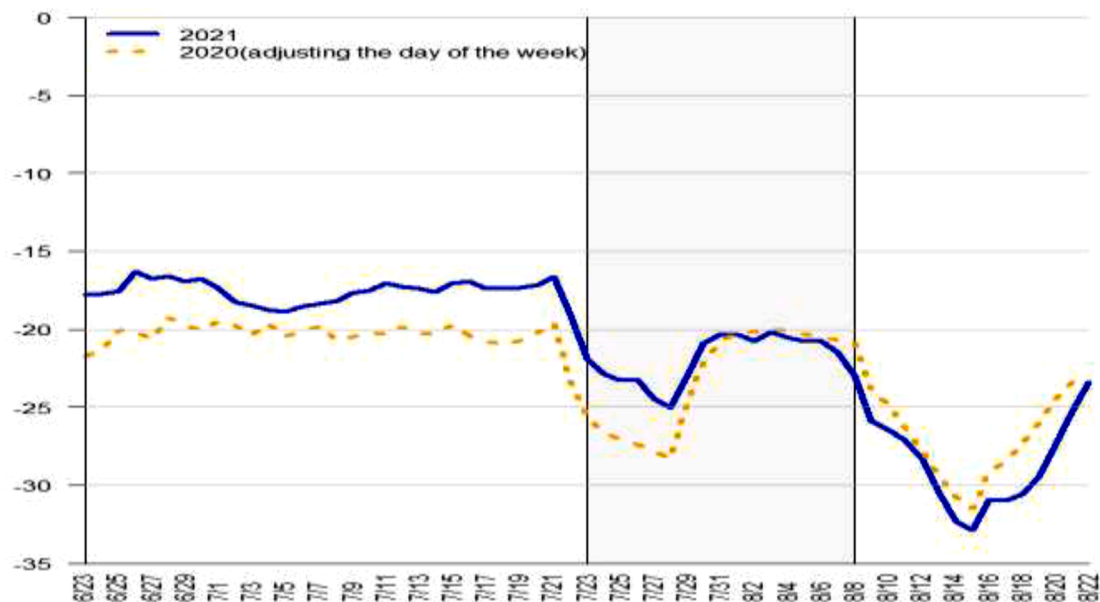
Appendix J. Mobility trends in Tokyo

Fig. J

Fig. J. Mobility trends in Tokyo: 2021 vs. 2020.
Notes: Fig. J shows a comparison of people's mobility in Tokyo in 2021 and 2020 during the same period, as in Fujii et al. (2021). We constructed the people's mobility index in Tokyo by using the data from the Google COVID-19 Community Mobility Reports. Following Chernozhukov et al. (2021), we focused on the four mobility measures: retail and recreation, groceries and pharmacies, transit stations, and work places. First, we calculated the simple average of the four mobility measures. Then, we calculated the 7-day moving average of the above value. To compare the mobility index in 2021 to that in 2020, the 2020 index was shifted by one day to match the day of the week in 2021.
Source: Our World in Data (https://ourworldindata.org/covid-google-mobility-trends). Accessed on August 28, 2021.

References

Abadie, Alberto., 2021. Using synthetic controls: feasibility, data requirements, and methodological aspects. J. Econ. Lit. 59 (2), 391–425.
Abadie, Alberto, Diamond, Alexis, Hainmueller, Jens, 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. J. Am. Statist. Assoc. 105 (490), 493–505.
Abadie, Alberto, Diamond, Alexis, Hainmueller, Jens, 2015. Comparative politics and the synthetic control method. Am. J. Polit. Sci. 59 (2), 495–510.
Abadie, Alberto, Gardeazabal, Javier, 2003. The economic costs of conflict: a case study of the basque country. Am. Econ. Rev. 93 (1), 113–132.
Altman, Douglas G., Bland, John Martin, 2011. How to obtain the confidence interval from a p value. The BMJ 343, d2090.
Asahi Shimbun. 2021a. "Safety and security of the Olympic Games "cannot be done" (68%)" (in Japanese). Poll conducted on July 17 and 18. Available online: https://www.asahi.com/articles/ASP7L7H2PP7HUZPS006.html.
Asahi Shimbun. 2021b. "61% of respondents said that the mood of self-restraint "had loosened" due to the Olympics" (in Japanese). Poll conducted on August 7 and 8. Available online: https://www.asahi.com/articles/ASP886DCNP86UZPS003.html.
Athey, Susan, Imbens, Guido W., 2017. The state of applied econometrics: causality and policy evaluation. J. Econ. Perspect. 31 (2), 3–32.
Ben-Michael, Eli, Feller, Avi, Rothstein, Jesse, 2021. The augmented synthetic control method. J. Am. Statist. Assoc. 116 (536), 1789–1803.
Botosaru, Irene, Ferman, Bruno, 2019. On the role of covariates in the synthetic control method. Econom. J. 22 (1), 117–130.
Bouttell, Janet, Craig, Peter, Lewsey, James, Robinson, Mark, Popham, Frank, 2018. Synthetic control methodology as a tool for evaluating population-level health interventions. J. Epidemiol. Commun. Health 72 (8), 673–678.
Breidenbach, Philipp, Mitze, Timo, 2022. Large-scale sport events and COVID-19 infection effects: evidence from the German professional football 'experiment'. Econom. J. 25 (1), 15–45.
Carlin, Patrick R., Minard, Paul, Simon, Daniel H., Wing, Coady, 2021. Effects of large gatherings on the COVID-19 epidemic: evidence from professional and college sports. Econ. Human Biol. 43, 101033.
Chernozhukov, Victor, Kasahara, Hiroyuki, Schrimpf, Paul, 2021. Causal impact of masks, policies, behavior on early Covid-19 pandemic in the US. J. Econom. 220 (1), 23–62.
Cho, Sang-Wook, 2020. Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: the case of Sweden. Econom. J. 23 (3), 323–344.

Dave, Dhaval, McNichols, Drew, Sabia, Joseph J., 2021. The contagion externality of a superspreading event: the Sturgis Motorcycle Rally and COVID-19. Southern Econ. J. 87 (3), 769–807.
Doudchenko, Nikolay, Imbens, Guido W., 2016. Balancing, regression, difference- in-differences and synthetic control methods: a synthesis. NBER Working Paper 22791.
Friedson, Andrew I., McNichols, Drew, Sabia, Joseph J., Dave, Dhaval, 2021. Shelter-in-place orders and public health: evidence from California during the COVID-19 pandemic. J. Policy Anal. Manage. 40 (1), 258–283.
Fujii, Daisuke, Nakata, Taisuke, 2021. COVID-19 and output in Japan. Japan. Econ. Rev. 72, 609–650.
Fujii, Daisuke, Maeda, Yuta, Mori, Masataka, Nagasawa, Kenichi, Nakata, Taisuke, Okamoto, Wataru, 2021. The effects of the Olympics and Paralympics on COVID-19 in Tokyo: ex-ante quantitative analyses. Mimeo.
Hayakawa, Kazunobu, Keola, Souknilanh, Urata, Shujiro, 2022. How effective was the restaurant restraining order against COVID-19? a nighttime light study in Japan. Japan World Econ. 63, 101136.
Hosono, Kaoru., 2021. Epidemic and economic consequences of voluntary and request-based lockdowns in Japan. J. Japan. Int. Econ. 61, 101147.
Kaul, Ashok, Klößner, Stefan, Pfeifer, Gregor, Schieler, Manuel, 2022. Standard synthetic control methods: the case of using all pre-intervention outcomes together with covariates. J. Bus. Econom. Statist. 40 (3), 1362–1376.
Kyodo News. 2021. "59% of respondents said that the Olympics were a factor in the spread of COVID-19" (in Japanese). Poll conducted on August 16. Available online: https://nordot.app/799894593932427264.
Linton, Natalie M., Sung-mok, Jung, Nishiura, Hiroshi, 2021. Not all fun and games: potential incidence of SARS-CoV-2 infections during the Tokyo 2020 Olympic Games. Math. Biosci. Eng. 18 (6), 9685–9696.
Mainichi Shimbun. 2021. "More than 70% believe that holding the Olympics would have affected the spread of COVID-19" (in Japanese). Poll conducted on August 28. Available online: https://mainichi.jp/articles/20210828/k00/00m/040/304000c.
Mitze, Timo, Kosfeld, Reinhold, Rode, Johannes, Wälde, Klaus, 2020. Face masks considerably reduce COVID-19 cases in Germany. Proc. Natl. Acad. Sci. 117 (51), 32293–32301.
Parnell, Daniel, Widdop, Paul, Bond, Alex, Wilson, Rob, 2020. COVID-19, networks and sport. Manag. Sport Leisure 1–7.
Ritchie, Hannah, Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian, and Max Roser. 2020. "Coronavirus pandemic (COVID-19)." Published online at OurWorldInData.org. Available online: https://ourworldindata.org/coronavirus.
Siordia, Jr, Juan, A, 2020. Epidemiology and clinical features of COVID-19: a review of current literature. J. Clin. Virol. 127, 104357.

Tokyo Organising Committee of the Olympic and Paralympic Games (TOCOG). 2021a. "Tokyo 2020 47th council information report" (in Japanese). Available online: https://www.tokyo2020.jp/image/upload/production/47EBMeeting6.pdf.

Tokyo Organising Committee of the Olympic and Paralympic Games (TOCOG). 2021b. "Results of COVID-19 countermeasures at the Tokyo 2020 Games." Available online: https://www.2020games.metro.tokyo.lg.jp/special/eng/guide/taikaijitorikumi/index.html.

Toumi, Asmae, Zhao, Haoruo, Chhatwal, Jagpreet, Linas, Benjamin P., Ayer, Turgay, 2021. Association of limited in-person attendance in US National Football League and National Collegiate Athletic Association games with county-level COVID-19 cases. JAMA Netw. Open 4 (8), e2119621.

Watanabe, Tsutomu, Yabu, Tomoyoshi, 2021a. Japan's voluntary lockdown. PLoS One 16 (6), e0252468.

Watanabe, Tsutomu, Yabu, Tomoyoshi, 2021b. Japan's voluntary lockdown: further evidence based on age-specific mobile location data. Japan. Econ. Rev. 72 (3), 333–370.