



OPEN

No causal effect of school closures in Japan on the spread of COVID-19 in spring 2020

Kentaro Fukumoto¹✉, Charles T. McClean^{2,3} and Kuninori Nakagawa⁴

Among tool kits to combat the coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2, school closures are one of the most frequent non-pharmaceutical interventions. However, school closures bring about substantial costs, such as learning loss. To date, studies have not reached a consensus about the effectiveness of these policies at mitigating community transmission, partly because they lack rigorous causal inference. Here we assess the causal effect of school closures in Japan on reducing the spread of COVID-19 in spring 2020. By matching each municipality with open schools to a municipality with closed schools that is the most similar in terms of potential confounders, we can estimate how many cases the municipality with open schools would have had if it had closed its schools. We do not find any evidence that school closures in Japan reduced the spread of COVID-19. Our null results suggest that policies on school closures should be reexamined given the potential negative consequences for children and parents.

School closures have been widely implemented as a non-pharmaceutical intervention (NPI) to reduce the spread of COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). By April 2020, 1 month after the World Health Organization (WHO) characterized COVID-19 as a pandemic, 173 countries had closed schools, affecting 84.3% of the world's enrolled students¹. Yet, school closures have broad impacts, including learning loss (as evidenced by the decrease in time spent learning² and test scores³), future earnings loss⁴, deterioration of physical health (for example, cancellation of school meals⁵ and increase in weight⁶) as well as mental health^{7,8}, maltreatment⁹ and lower maternal labor force participation¹⁰, including health-care workers¹¹. Notably, these costs are disproportionately imposed on disadvantaged families, thereby widening social and economic inequality^{2–4}. Furthermore, school closures will lead to even long-term macroeconomic damage^{4,12}.

Accordingly, it is imperative to know whether the benefits of school closures outweigh these costs. Nonetheless, scholars have not reached a consensus on the degree of benefit, if any, to closing (or not reopening) schools (see a recent systematic review¹³). Some simulation^{14–19} and empirical^{20–33} studies show that school closures are effective in mitigating the spread of COVID-19. However, others fail to establish such statistically significant effects^{33–49}.

We argue that one of reasons why the literature is equivocal is methodological. Simulation studies assume parameters in their models, whose values may not be correct. Most empirical works estimate parameters (including the effect of school closures) by using publicly available aggregated data, although these studies are not necessarily rigorous in terms of causal inference. A typical research design is panel regression: using a dataset that spans across countries and days, researchers regress the number of cases on a dummy variable to indicate whether a country closes its schools on a given day, where the coefficient of the dummy represents the effect of school closures. In essence, researchers estimate the effect of school closures by measuring the difference in the number of cases between days when a country closes and opens its schools.

Many articles do not control for any other variables, while others include only a few control variables. Therefore, readers should be concerned about dozens of potential confounders that affect both school closures and the number of cases (for example, share of children in the population, medical preparedness and the government's fiscal situation), which would bias estimates of the effect of school closures. Relatedly, we cannot rule out the possibility of reverse causality, namely, that governments close their schools exactly because of a high number of COVID-19 cases among their residents⁴². If this is true, naive regressions with few control variables would underestimate the effect of school closures on infection. Moreover, in essence, panel regression exploits variation in school closures across space and time, but school closures usually coincide with other NPIs (such as stay-at-home orders and prohibitions on gatherings) and/or are introduced simultaneously nationwide. Thus, it is challenging to disentangle the effect of school closures from those of other NPIs and/or from other contemporary factors such as season, economy and weather, especially when the unit of analysis is as large as a country or state^{13,27,32,35}.

Here, to estimate the causal effect of school closures on reducing the spread of COVID-19, we use data from Japan, where some municipalities closed their schools, while others did not. We exploit this variation in school closures among hundreds of municipalities by utilizing matching techniques to account for dozens of confounders⁵⁰ (Table 1).

Results

Data. In our study, the unit of observation is a municipality. In Japan, each municipality is responsible for the closures of its elementary (K1–6, students aged 6 to 12 years) and junior high (K7–9, students aged 12 to 15 years) schools (on average, 13.5 and 6.9 schools per municipality, respectively). In principle, children who attend public schools do so in the same municipality in which they reside. School closures began on 2 March 2020. The Ministry of Education, Culture, Sports, Science and Technology (MEXT) conducted municipal-level surveys on whether schools were open eight

¹Department of Political Science, Gakushuin University, Tokyo, Japan. ²Program on U.S.-Japan Relations, Harvard University, Cambridge, MA, USA.

³Center for Japanese Studies, University of Michigan, Ann Arbor, MI, USA. ⁴Department of Economics, Shizuoka University, Shizuoka, Japan.

✉e-mail: kentaro.fukumoto@gakushuin.ac.jp

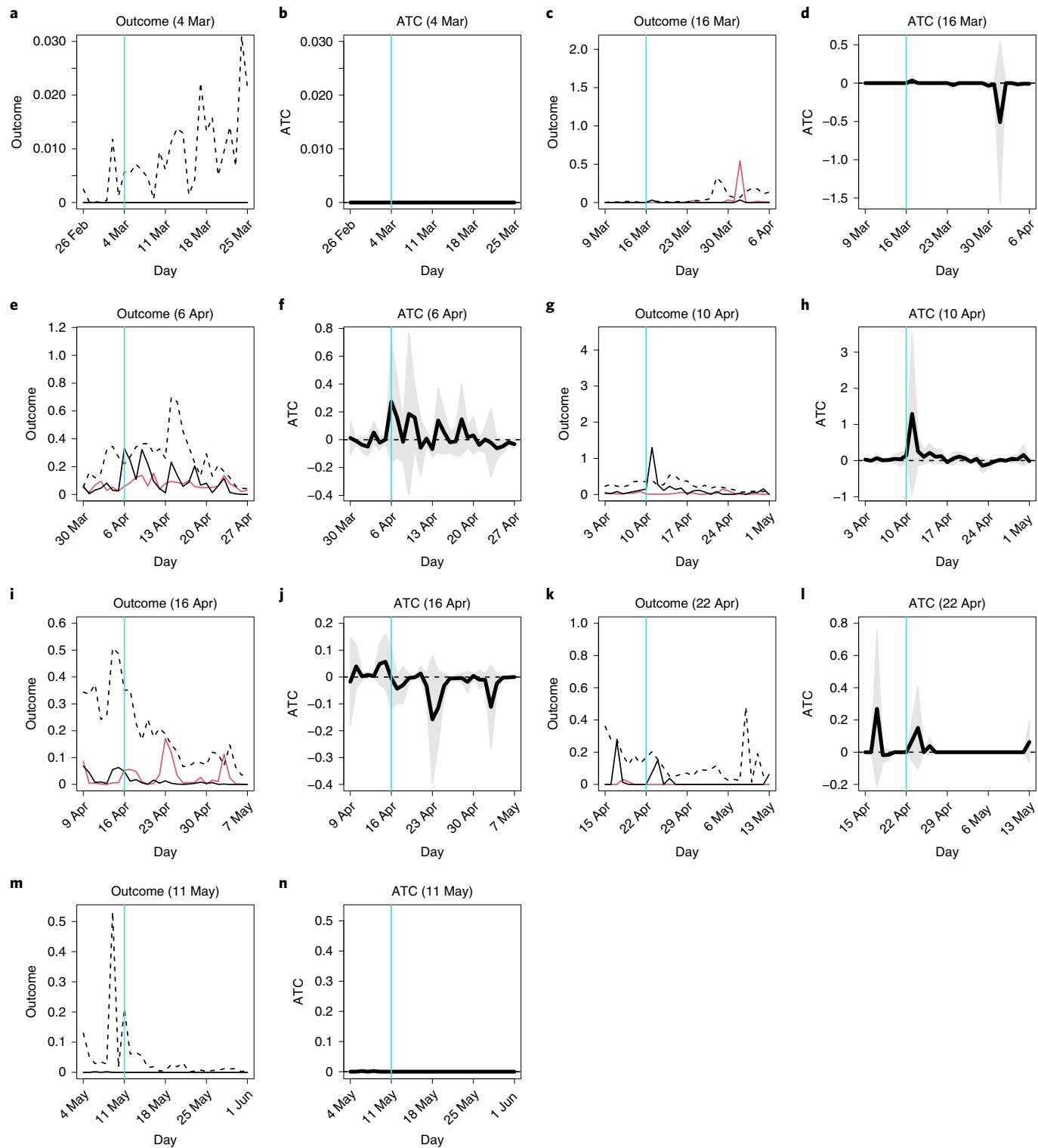


Fig. 1 | Average outcomes and ATC values: main analysis. **a–n**, The horizontal axis indicates dates in 2020, and the vertical turquoise line corresponds to the survey date. In **a, c, e, g, i, k** and **m**, the vertical axis represents the average number of confirmed cases per 100,000 residents, the black and red lines correspond to the average outcomes of the matched treated and control municipalities, respectively, and the dashed black line represents the average outcomes of all treated municipalities, both matched and unmatched. In **b, d, f, h, j, l** and **n**, the vertical axis represents ATC values, the thick black line indicates the point estimates of ATC values, and the shaded gray area presents the 95% confidence intervals. Cluster-robust standard errors were utilized where clusters were a pair of matched municipalities and a municipality. **a**, Outcome, 4 March. **b**, ATC, 4 March. **c**, Outcome, 16 March. **d**, ATC, 16 March. **e**, Outcome, 6 April. **f**, ATC, 6 April. **g**, Outcome, 10 April. **h**, ATC, 10 April. **i**, Outcome, 16 April. **j**, ATC, 16 April. **k**, Outcome, 22 April. **l**, ATC, 22 April. **m**, Outcome, 11 May. **n**, ATC, 11 May.

Table 1 | Policy summary**Background**

Studies have not reached a consensus about the effectiveness of school closures on reducing the spread of COVID-19, partly because of analytical challenges. To estimate the causal effects strictly in Japan in spring 2020, we compared the number of cases between municipalities with open schools and municipalities with closed schools, both of which were selected so that they are similar in terms of dozens of features (for example, past number of cases; social, economic and political factors; and NPIs other than school closures such as stay-at-home warnings and business closures).

Main findings and limitations

On average, the numbers of cases per 100,000 residents in municipalities with closed schools were not statistically significantly smaller than those in municipalities with open schools. Because both groups are comparable except for school closure status, the difference between them should only be attributed to school closure status. Thus, we did not find any evidence that school closures significantly decreased the spread of COVID-19. It is possible that school closures lowered COVID-19 cases outside Japan and/or after spring 2020, in particular in the presence of new variants of SARS-CoV-2.

Policy implications

Our null results concerning the supposed benefits of closing schools suggest that policymakers should be cautious when considering similar policies in the future, especially given the substantial costs to the well-being of both children and parents. Our recommendation is that governments should monitor SARS-CoV-2 infection rates and school closures at a granular level (for example, municipality or school district) in real time (for example, daily) to inform their policy decisions.

times in 2020: 4 and 16 March; 6, 10, 16 and 22 April; 11 May; and 1 June (Methods). For each survey date, the treatment variable is equal to one if all elementary and junior high schools in the municipality are closed as of the survey date and zero if they are open. When some schools are closed but others are open, the treatment variable has a missing value.

The outcome variables are the daily numbers of newly confirmed cases of COVID-19 per 100,000 residents of the municipality. Due to the availability of the outcome variables, we only use data from 26 (for 6 April and earlier treatments) or 27 (for 10 April and later treatments) of the 47 prefectures in Japan (Methods and Extended Data Fig. 1).

To address confounding, we controlled for dozens of covariates (Methods). First, we controlled for school closure status as of each past survey date, the total number of COVID-19 cases before the survey date and the number of COVID-19 cases in each of the past 7 d. We also included a set of 25 or 26 prefecture dummy variables, each of which indicates whether a given municipality belongs to the corresponding prefecture, to account for variation across prefectures. In Japan, prefectures have the primary responsibility for public health, including infectious disease. The governor of a prefecture can issue (de jure or de facto) NPI requests such as stay-at-home warnings, business closures and event suspensions. Because these NPIs are requests, not orders, citizens have no legal obligation to obey such NPIs. Thus, by including prefecture dummies, we could account for any characteristics of prefectures, including NPIs, which differed across prefectures during the time of our study. As a result, there was no remaining variation in NPIs among municipalities other than school closures. For this reason, we believe that our research design can identify the effect of school closures. We also controlled for latitude and longitude so that closer municipalities are more likely to be matched. Additionally, we controlled for 39 social, economic and political variables (Methods and Extended Data Fig. 2).

Table 2 | Number of treated (closed schools) and control (open schools) municipalities by survey date

Survey date	Treated	Control	All	Treated (%)
4 March	761	10	771	98.7
16 March	718	29	747	96.1
6 April	256	483	739	34.6
10 April	491	307	798	61.5
16 April	523	267	790	66.2
22 April	710	80	790	89.9
11 May	641	145	786	81.6
1 June	2	783	785	0.3

There are 1,741 municipalities in Japan. This number was reduced to 785 or 847 when we limited our analysis to the 26 or 27 prefectures mentioned above. Moreover, if any variable had a missing value in an analysis for a treatment variable, that municipality was not used in the analysis for the treatment variable. For details of our variables, see Methods. Table 2 presents the number of treated and control municipalities by survey date. Below, our analysis covers up to the 11 May survey date (Methods).

Matching. It is challenging to estimate causal effects from observational data rather than experimental data mainly because it is hard to control for confounders. Matching is one commonly used method to address these problems^{27,50}. Here, for every control municipality, we matched a treated municipality that had similar values across our covariates to those of the control municipality (Methods). The average of the matched treated municipalities' outcomes can identify the average of the counterfactual outcomes for the control municipalities if they were assigned treatment. Therefore, by subtracting the average of the control municipalities from that of the matched treated municipalities, we can estimate the average treatment effect on the control (ATC)⁵⁰, namely, the degree to which school closures would change the number of COVID-19 cases in municipalities that actually did not close their schools (Methods). That is, if school closures have a causal effect to reduce the number of cases, the ATC should be negative. Note that most empirical studies (implicitly) assume some type of model (such as a linear model), while our identification strategy does not.

Recall that we do not know whether municipalities closed or opened their schools between the survey dates. In addition, most previous work focused on 7 to 14 d after implementation of an NPI, considering periods of incubation, test and report delay¹³. Accordingly, we cannot conduct panel data analysis from March to June 2020. Instead, for every set of survey date and outcome date, we implemented a cross-sectional analysis across municipalities where we performed matching and regressed the outcome variable on the treatment variable using only the matched municipalities.

Main analysis. Figure 1 presents the results. Each panel corresponds to either the average outcomes (number of COVID-19 cases per 100,000 municipal residents) or ATC values for the treatment variable (school closure status) as of a survey date. As an example, we will focus on the 6 April treatment variable (Supplementary Information; Main analysis, 6 April). In Fig. 1e, the horizontal axis indicates the dates of outcome variables, and the vertical axis represents the values of outcome variables. We examine 7 d before the survey date to 21 d after the survey date. The vertical turquoise line marks the survey date. The red line corresponds to the average outcome values of the control municipalities. The dashed black line shows the average outcome values of all treated municipalities, both

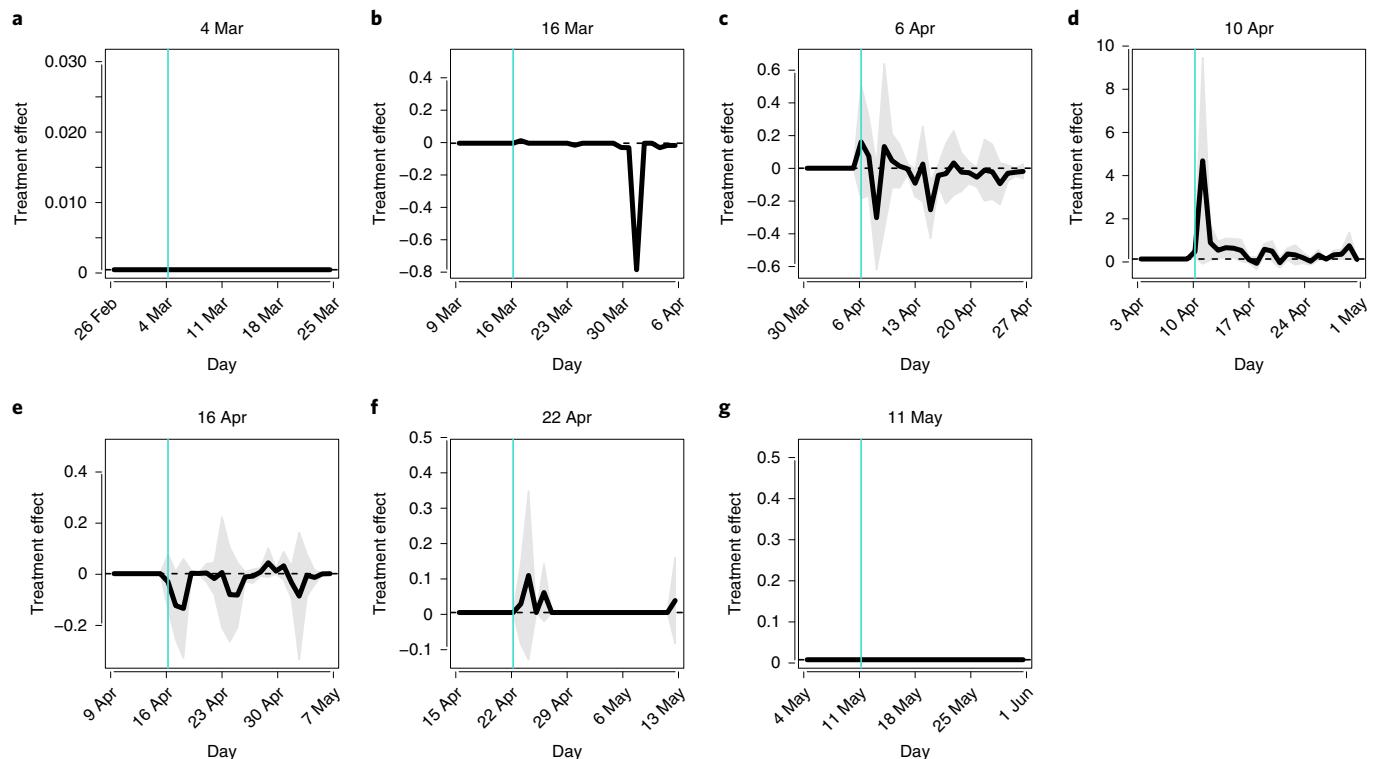


Fig. 2 | Treatment effects: controlling for covariates. a–g. The horizontal axis indicates dates in 2020, and the vertical axis represents the coefficient estimates of the treatment variable. The vertical turquoise line corresponds to the survey date. The thick black line indicates the point estimates of the treatment variable's coefficients, and the shaded gray area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters were a pair of matched municipalities and a municipality. **a**, 4 March. We did not draw the shaded gray area because the estimation procedure did not return the standard errors. **b**, 16 March. **c**, 6 April. **d**, 10 April. **e**, 16 April. **f**, 22 April. **g**, 11 May.

matched and unmatched. If we compare the number of COVID-19 cases between the control (red line) and treated (dashed black line) municipalities without taking matching into account, the latter group has a higher number of cases than the former every day. But these two groups are not comparable. When we pay attention to the period before the survey date, the average number of cases in both groups follows distinct trajectories. This may suggest that control municipalities kept their schools open exactly because, in consideration of school closures, they had fewer cases, and thus they were more confident that they would be free from COVID-19 than treated municipalities.

To address concerns such as confounders and/or reverse causality, we shift to the solid black line, which indicates the average outcome values of the matched treated municipalities. Now, because both treated and control group averages almost overlap before the survey date, we can see that the matched groups share a similar infection history. This test is similar to the parallel trend check in the differences-in-differences technique⁴⁶. Moreover, the differences in other covariates between the treated and control groups were also much smaller after matching than before (Supplementary Fig. 1 and Supplementary Table 3). Therefore, differences between the matched groups cannot be attributed to previous levels of infection or any other covariates.

Even after the survey date, the average outcomes of the matched treated municipalities are for the most part not smaller than those of the control municipalities, including 1–2 weeks after the survey date, the period that is the focus of most studies. But recall that the matched treated municipalities closed their schools even though they were in a similar situation to the control municipalities. If school closures have an effect to reduce the spread of COVID-19, the matched treated municipalities would have fewer cases than the

control municipalities (or the control municipalities would fail to prevent an outbreak and have more cases). Instead, we find few significant differences in case numbers between the matched treated and control municipalities.

Correspondingly, the ATC values are almost null. In Fig. 1f, the vertical axis represents the estimates of the ATC values, where the thick black line indicates the point estimates, and the shaded gray area presents the 95% confidence intervals (Methods). Note that by subtracting the average outcome values of the control municipalities (red line in Fig. 1e) from those of the matched treated municipalities (black line in Fig. 1e), we obtained the ATC point estimates (thick black line in Fig. 1f). The ATC values were rarely significantly negative. Therefore, we cannot say that the school closures treatment as of 6 April had a causal effect on reducing the number of cases.

Other panels of Fig. 1 show the average outcomes and ATC values of other treatment variables in a similar fashion (Supplementary Information; main analysis). The implications of our findings for 6 April also hold for the remaining survey dates: school closures did not significantly reduce the spread of COVID-19 in Japan between 4 March and 1 June 2020. For some irregular spikes of ATC values, see the Supplementary Information (main analysis, 16 March)

Robustness checks. To confirm that the above results are robust, we conducted a battery of robustness checks.

Controlling for covariates. By matching, we reduced the differences in covariates between treated and control groups, although we failed to remove them altogether (Supplementary Fig. 1). Following recommendations in ref. ⁵¹, we regressed the outcomes not only on a treatment but also on the covariates that we matched on to further

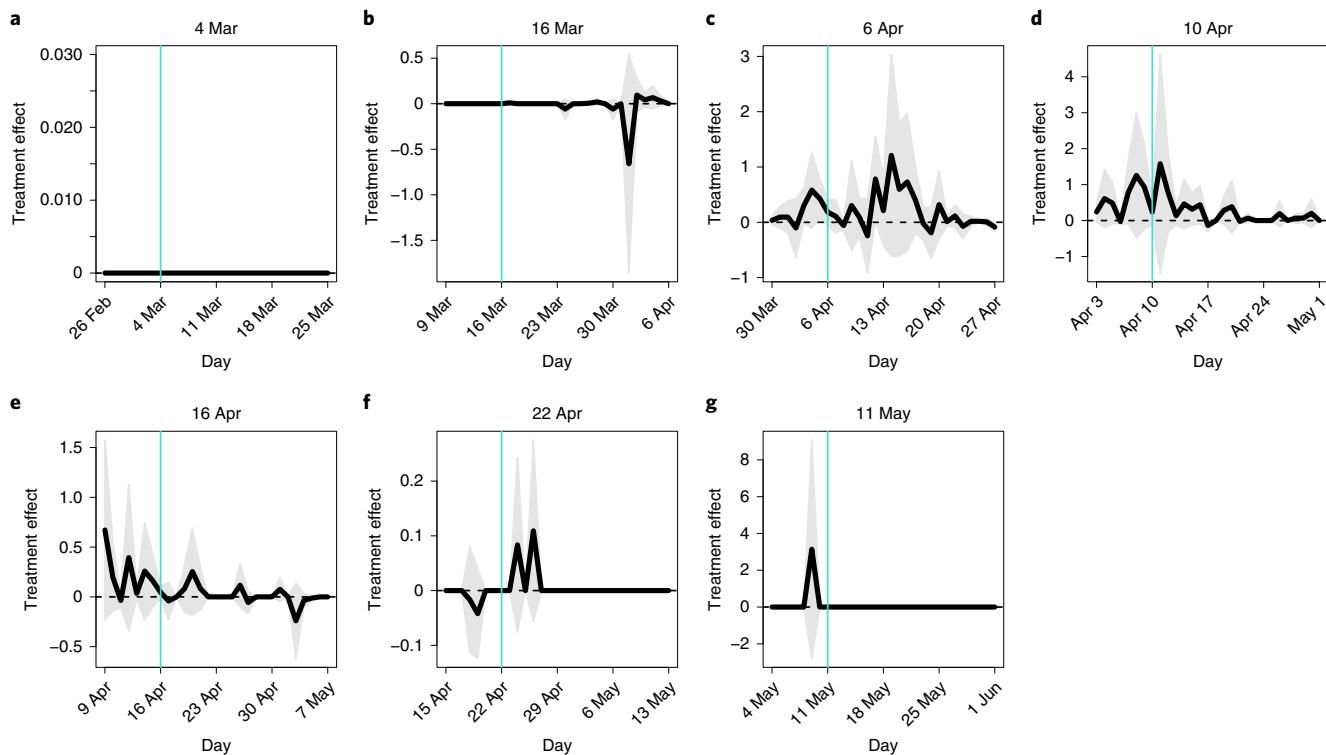


Fig. 3 | Treatment effects: public health center fixed effects. **a–g.** The horizontal axis indicates dates in 2020, and the vertical axis represents the coefficient estimates of the treatment variable. The vertical turquoise line corresponds to the survey date. The thick black line indicates the point estimates of the treatment variable's coefficients, and the shaded gray area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters were a public health center. **a.** 4 March. **b.** 16 March. **c.** 6 April. **d.** 10 April. **e.** 16 April. **f.** 22 April. **g.** 11 May.

decrease the effects of confounders. Each panel of Fig. 2 shows the coefficient estimates of the corresponding treatment variable. The ATC values again suggest that school closures do not lower the infection of SARS-CoV-2 significantly.

Negative binomial regression. Because the numbers of cases are non-negative integers (Supplementary Fig. 2), some studies rely on negative binomial regression^{21,22,41}. Therefore, we also applied negative binomial regression to our matched data⁵², using the log of the population size as an offset variable so that various sized municipalities are comparable. Unfortunately, the estimation procedures only converged for the treatment variables as of 6, 10 and 16 April, and only if we did not include any covariates. The coefficient estimates of the treatment variables sometimes implied that school closures even significantly increased the number of cases, while they rarely (significantly) decreased the number of cases (Extended Data Fig. 3). The findings remain the same: there is no evidence that school closures reduce the number of COVID-19 cases.

Public health center fixed effects. Public health centers in Japan are in charge of implementing public health policy, including counting COVID-19 cases. There are a total of 263 public health centers in the 27 prefectures in this study, each covering 1 to 13 municipalities (Supplementary Fig. 3). To address heterogeneity in the centers' efforts at confirming and/or tracking cases, we regressed each of our outcome variables on one of our treatment variables and public health center dummy variables, using both matched and unmatched municipalities. We identified the effect of the treatment variable by (a weighted average of) the difference in means of the outcome variable between treated and control municipalities within a public health center (Supplementary Table 5). Because a public health center covers adjacent municipalities, its fixed effect controls for any

idiosyncratic factors not only held by the public health center itself but also shared by the municipalities it covers (for example, urbanicity), effectively substituting for many covariates we did not technically control for. Figure 3 illustrates the coefficient estimates of the treatment variables. Again, we did not find in this analysis that school closures have significant effects on the number of COVID-19 cases.

Inverse probability weighting. An alternative to matching is inverse probability weighting²⁰ (Methods). When we estimated the 'probability' (that is, propensity score and thus weight), we used the same covariates as in matching except latitude and longitude. By subtracting the average outcome of the control municipalities from the weighted average of the treated municipalities (Extended Data Fig. 4a,c,e,g,i,k,m), we could estimate the ATC (Extended Data Fig. 4b,d,f,h,j,l,n). The ATC values suggest that municipalities that closed their schools mostly increased the number of cases (Supplementary Information; robustness checks, inverse probability weighting) and did not significantly decrease the spread of COVID-19 for the vast majority of dates. The results further support our main findings.

Conditioning on neighbors. Currently, most causal inference studies assume that the outcome of a municipality is independent of other municipalities' treatments, which is implied by the stable unit treatment value assumption⁵⁰. But in our context, the treatment and outcome in one municipality will affect those in another⁴⁶. It is challenging to take into consideration such spatial spillover.

One state-of-the-art study proposes the stable unit treatment on neighborhood value assumption: the outcome of a municipality is independent of other municipalities' treatments conditioned on (the average of) the treatments of neighboring municipalities⁵³. For all of the survey dates except 6 April, the most frequent pattern was the situation where a municipality was surrounded by only treated

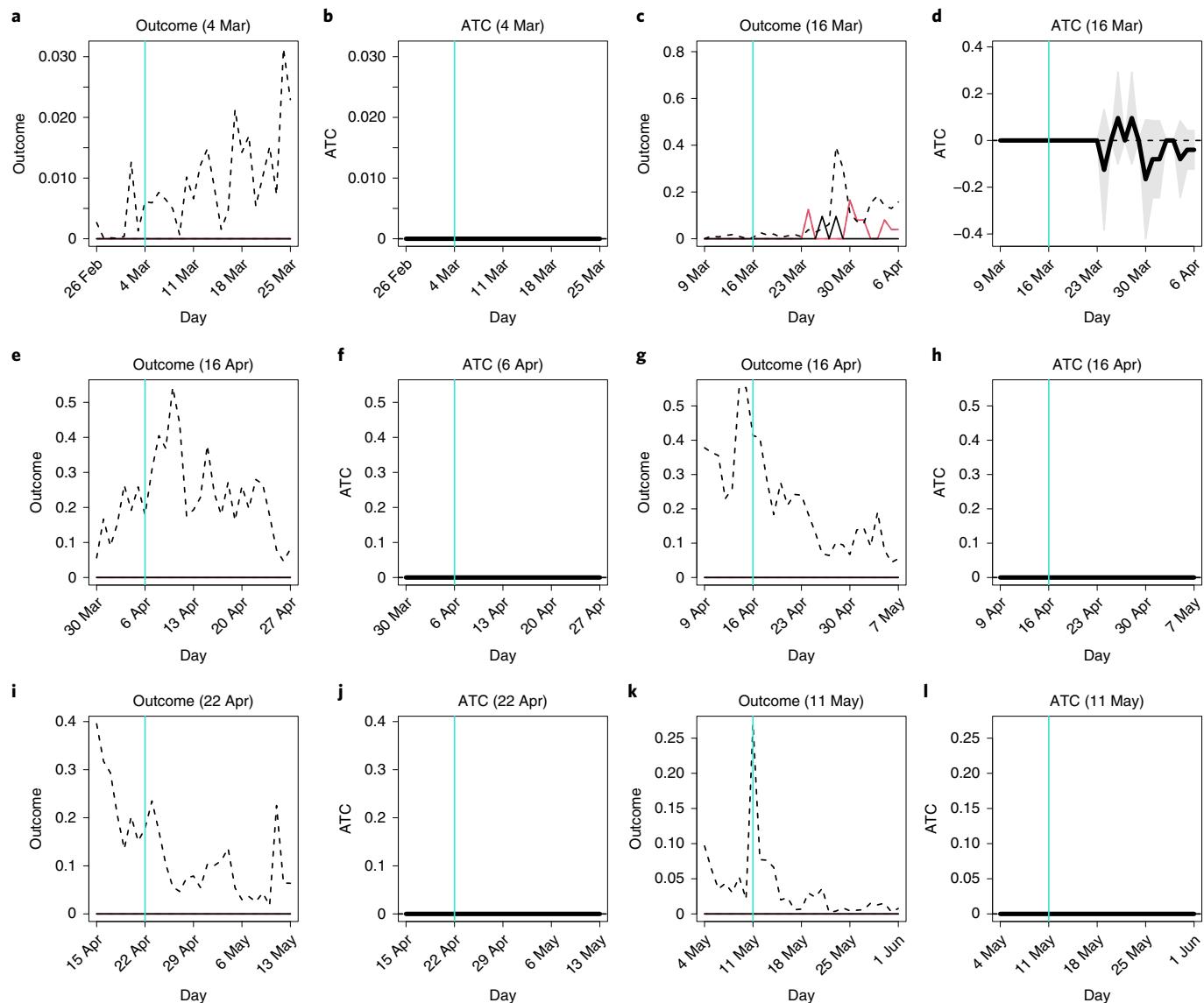


Fig. 4 | Average outcomes and ATC values: conditioning on neighbors. **a–k**, the horizontal axis indicates dates in 2020, and the vertical turquoise line corresponds to the survey date. In **a, c, e, g, i** and **k**, the vertical axis represents the average number of confirmed cases per 100,000 residents, the black and red lines correspond to the average outcomes of the matched treated and control municipalities, respectively, and the dashed black line represents the average outcomes of all treated municipalities, both matched and unmatched. In **b, d, f, h, j** and **l**, the vertical axis represents ATC values, the thick black line indicates the point estimates of ATC values, and the shaded gray area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters were a pair of matched municipalities and a municipality. **a**, outcome, 4 March. **b**, ATC, 4 March. **c**, Outcome, 16 March. **d**, ATC, 16 March. **e**, Outcome, 6 April. **f**, ATC, 6 April. **g**, Outcome, 16 April. **h**, ATC, 16 April. **i**, Outcome, 22 April. **j**, ATC, 22 April. **k**, Outcome, 11 May. **l**, ATC, 11 May.

municipalities (on 6 April, this situation was the second most frequent; Supplementary Table 6 and Supplementary Fig. 4); therefore, we estimated the ATC value given all neighbor municipalities are treated by limiting our analyses to only those municipalities whose neighboring municipalities were all treated. We applied matching to those municipalities. Because this condition is restrictive, we only ended up with two to six pairs of matched municipalities for each survey date (Supplementary Table 7). Thus, we warn that the results may not be reliable. For the treatment on 10 April, there remained only one control municipality, and the matching algorithm failed to match any treated municipality to it; therefore, we did not report the corresponding results.

As shown in Fig. 4, for most days, treated and control municipalities had no cases. Therefore, the ATC values imply that school closures do not change the number of cases, although they either

increase or decrease the number of cases, albeit insignificantly, on several days. If opening schools leads to the spread of COVID-19, spikes of cases would occur in the control group; however, these were not observed. The implication is the same: school closures do not help reduce the spread of COVID-19 significantly.

Average treatment effect on the treated. We next estimated the average treatment effect on the treated (ATT), namely, the degree to which school closures changed the number of COVID-19 cases in municipalities that actually did close their schools compared to how many cases those same municipalities would have if their schools were instead open (Methods). To estimate an ATT, we compare a sufficiently similar control municipality to every treated municipality. Because only the 6 April survey date had more control municipalities than treated ones (Table 2), we only estimated the ATT values for

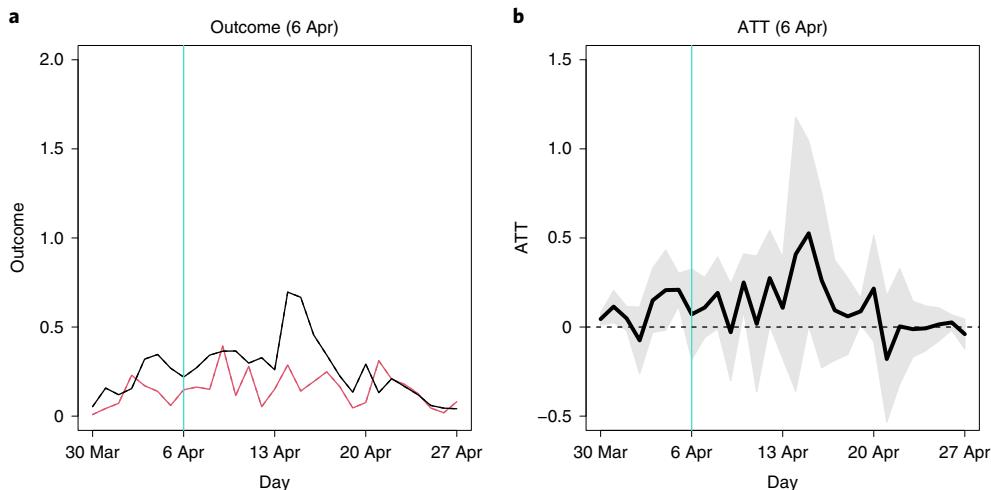


Fig. 5 | Average outcomes and ATT values. **a,b,** The horizontal axis indicates dates in 2020. The vertical turquoise line corresponds to the survey date, 6 April. **a,** The vertical axis represents the average number of confirmed cases per 100,000 residents. The black and red lines correspond to the average outcomes of the matched treated and control municipalities, respectively. **b,** The vertical axis represents ATT values. The thick black line indicates the point estimates of ATTs, and the shaded gray area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters were a pair of matched municipalities and a municipality.

the 6 April treatment variable. Compared with the matched control municipalities (red line in Fig. 5a), the (matched) treated municipalities (black line in Fig. 5a) had similar average outcomes before the survey date and mostly larger average outcomes after the survey date. (In Fig. 5a, the dashed black line (all of the treated municipalities) overlaps the solid black line and is thus invisible because all of the treated municipalities are matched to control municipalities.) Accordingly, the ATT values never indicate that school closures significantly reduced local COVID-19 cases (Fig. 5b). Therefore, our null findings apply not only for ATC values but also ATT values, at least for the 6 April treatment.

A smaller set of covariates. Because we matched on dozens of covariates, readers may be concerned about potential over-adjusting for confounding and collider bias. To address this concern, in the setup of the main analysis, we matched on only 16 of the 39 social, economic and political variables in addition to the past treatment variables, the set of prefecture dummies and eight outcomes (the total number of COVID-19 cases before the survey date and the number of COVID-19 cases in each of the past 7 d; Methods). Again, our results fail to confirm that school closures decrease the number of COVID-19 cases (Extended Data Fig. 5).

Discussion

We identify the effect of school closures on the spread of COVID-19 using matching methods. Our results suggest that the effect of school closures on COVID-19 cases in Japan in early 2020 is not significantly different from zero. Ref. ⁴⁵ comes to the same conclusion based on different data from Japan and a different method.

Our research design does not enable us to elucidate the mechanisms underlying the empirical findings. That said, we provide a few conjectures. First, many treated municipalities did not comply with the treatment in the sense that many schools that were closed nevertheless provided spaces for children whose parents worked such as playing fields, gyms, classrooms and libraries. For instance, schools made these facilities available to families in 63.6% of the treated municipalities as of 17 March and 59% as of 16 April (Supplementary Information; main analysis, 16 March and 16 April). Furthermore, children may interact with each other outside of school^{19,26,36,41}. Second, as past studies have demonstrated^{19,36,40,54},

even in control municipalities, children may be less susceptible to SARS-CoV-2 and less likely to transmit the virus to others, including teachers, parents and neighbors. Moreover, students attending in-person schooling were engaged in stringent mitigation strategies such as physical distancing and enhanced hygiene. Limiting the amount of contact between children by shutting down schools may thus have not had a large impact on reducing the spread of COVID-19.

We admit some limitations of our study. Although we are confident in the internal validity of our findings, we are less sure about their external validity (Supplementary Information). Japan, like New Zealand and South Korea, fared relatively well with regard to COVID-19 cases compared to some countries such as Italy, Spain and the United States. It is possible that school closures only have a discernible effect on COVID-19 cases once the reproduction number (within a school environment) passes a certain threshold. Likewise, we analyze the effectiveness of school closures in the initial months of the coronavirus pandemic, but community transmission has become notably worse in many countries (including Japan) since spring 2020. New variants of SARS-CoV-2 (for example, the alpha and delta variants), which have each increased transmissibility, may also be more or less affected by school closures. Another possibility is that citizens' behavior (for example, social distancing, wearing masks and washing hands) and school measures (for example, disinfecting classrooms, placing transparent dividers between students, splitting classes into smaller groups that meet separately and ventilation) intended to limit the spread of COVID-19 may be less strict in some countries compared to Japan, which could increase the risk of opening schools.

With these caveats in mind, we offer both empirical and methodological contributions to the existing literature. Empirically, we find no evidence that school closures in Japan caused a significant reduction in the number of coronavirus cases. These null results concerning the supposed benefits of closing schools suggest that policymakers should be cautious when considering similar policies in the future, especially given the substantial costs such policies can have for the well-being of both children and parents. Our recommendation is that governments should monitor SARS-CoV-2 infection rates and school closures at a granular level (for example, municipality or school district) in real time (for example, daily).

Methodologically, we pay attention to causal inference by using matching techniques and exploiting features of Japanese municipalities that enable us to identify the effect of school closures independent of other NPIs. Our hope is that our study can inspire other researchers to use causal inference methods to rigorously test the generalizability of our findings to other settings.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-021-01571-8>.

Received: 24 July 2021; Accepted: 1 October 2021;

Published online: 27 October 2021

References

- School closures caused by coronavirus (COVID-19)* (UNESCO, 2020); <https://en.unesco.org/covid19/educationresponse> (accessed 27 January 2021).
- Andrew, A. et al. Inequalities in children's experiences of home learning during the COVID-19 lockdown in England. *Fisc. Stud.* **41**, 653–683 (2020).
- Engzell, P., Frey, A. & Verhagen, M. D. Learning loss due to school closures during the COVID-19 pandemic. *Proc. Natl Acad. Sci. USA* **118**, e20223761118 (2021).
- Hanushek, E. A. & Woessmann, L. The economic impacts of learning losses. Preprint at <https://www.oecd-ilibrary.org/content/paper/21908d74-en> (2020).
- Kinsey, E. W. et al. School closures during COVID-19: opportunities for innovation in meal service. *Am. J. Public Health* **110**, 1635–1643 (2020).
- Takaku, R. & Yokoyama, I. What the COVID-19 school closure left in its wake: evidence from a regression discontinuity analysis in Japan. *J. Public Econ.* **195**, 104364 (2021).
- Yamamura, E. & Tsutsui, Y. School closures and mental health during the COVID-19 pandemic in Japan. *J. Popul. Econ.* **34**, 1261–1298 (2021).
- Xie, X. et al. Mental health status among children in home confinement during the coronavirus disease 2019 outbreak in Hubei Province, China. *JAMA Pediatr.* **174**, 898–900 (2020).
- Baron, E. J., Goldstein, E. G. & Wallace, C. T. Suffering in silence: how COVID-19 school closures inhibit the reporting of child maltreatment. *J. Public Econ.* **190**, 104258 (2020).
- Collins, C., Ruppanner, L., Landivar, L. C. & Scarborough, W. J. The gendered consequences of a weak infrastructure of care: school reopening plans and parents' employment during the COVID-19 pandemic. *Gend. Soc.* **35**, 180–193 (2021).
- Bayham, J. & Fenichel, E. P. Impact of school closures for COVID-19 on the US health-care workforce and net mortality: a modelling study. *Lancet Public Health* **5**, e271–e278 (2020).
- Jang, Y. & Yum, M. Aggregate and intergenerational implications of school closures: a quantitative assessment. *Covid Econ.* **57**, 46–93 (2020).
- Walsh, S. et al. Do school closures and school reopening affect community transmission of COVID-19? A systematic review of observational studies. *BMJ Open* **11**, e053371 (2021).
- Castex, G., Dechter, E. & Lorca, M. COVID-19: the impact of social distancing policies, cross-country analysis. *Econ. Disaster Clim. Chang.* **5**, 135–159 (2021).
- Chen, X. & Qiu, Z. Scenario analysis of non-pharmaceutical interventions on global COVID-19 transmissions. *Covid Econ.* **7**, 50–71 (2020).
- Head, J. R. et al. School closures reduced social mixing of children during COVID-19 with implications for transmission risk and school reopening policies. *J. R. Soc. Interface* **18**, 20200970 (2021).
- Kim, S., Kim, Y.-J., Peck, K. R. & Jung, E. School opening delay effect on transmission dynamics of coronavirus disease 2019 in Korea: based on mathematical modeling and simulation study. *J. Korean Med. Sci.* **35**, 57–61 (2020).
- Panovska-Griffiths, J. et al. Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. *Lancet Child Adolesc. Health* **4**, 817–827 (2020).
- Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).
- Amadio, E., Battisti, M., Kourtellos, A., Maggio, G. & Maida, C. M. Schools opening and COVID-19 diffusion: evidence from geolocalized microdata. *Covid Econ.* **65**, 47–77 (2021).
- Auger, K. A. et al. Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA* **324**, 859–870 (2020).
- Banholzer, N. et al. Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLoS ONE* **16**, e0252827 (2021).
- Brauner, J. M. et al. Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, eabd9338 (2021).
- Chernozhukov, V., Kasahara, H. & Schrimpf, P. Causal impact of masks, policies, behavior on early COVID-19 pandemic in the U.S. *J. Econom.* **220**, 23–62 (2021).
- Chernozhukov, V., Kasahara, H. & Schrimpf, P. The association of opening K-12 schools and colleges with the spread of COVID-19 in the United States: county-level panel data analysis. Preprint at <https://www.medrxiv.org/content/early/2021/06/16/2021.02.20.21252131> (2021).
- Courtemanche, C. J., Le, A. H., Yelowitz, A. & Zimmer, R. School reopenings, mobility, and COVID-19 spread: evidence from Texas. Preprint at <http://www.nber.org/papers/w28753> (2021).
- Haug, N. et al. Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat. Hum. Behav.* **4**, 1303–1312 (2020).
- Li, Y. et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *Lancet Infect. Dis.* **21**, 193–202 (2021).
- Liu, Y. et al. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BMC Med.* **19**, 40 (2021).
- Matzinger, P. & Skinner, J. Strong impact of closing schools, closing bars and wearing masks during the COVID-19 pandemic: results from a simple and revealing analysis. Preprint at <https://www.medrxiv.org/content/early/2020/09/28/2020.09.26.20202457> (2020).
- Stage, H. B. et al. Shut and re-open: the role of schools in the spread of COVID-19 in. *Eur. Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200277 (2021).
- Vlachos, J., Hertegård, E. & Svaleryd, H. B. The effects of school closures on SARS-CoV-2 among parents and teachers. *Proc. Natl Acad. Sci. USA* **118**, e2020834118 (2021).
- Yehya, N., Venkataramani, A. & Harhay, M. O. Statewide interventions and coronavirus disease 2019 mortality in the United States: an observational study. *Clin. Infect. Dis.* **73**, 1863–1869 (2020).
- Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M. & Prokopenko, M. Modelling transmission and control of the COVID-19 pandemic in Australia. *Nat. Commun.* **11**, 5710 (2020).
- Davies, N. G. et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *Lancet Public Health* **5**, e375–e385 (2020).
- Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).
- Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
- Lee, B., Hanley, J. P., Nowak, S., Bates, J. H. T. & Hébert-Dufresne, L. Modeling the impact of school reopening on SARS-CoV-2 transmission using contact structure data from Shanghai. *BMC Public Health* **20**, 1713 (2020).
- Wang, X. et al. Impact of social distancing measures on coronavirus disease healthcare demand, central Texas, USA. *Emerg. Infect. Dis.* **26**, 2361–2369 (2020).
- Gandini, S. et al. A cross-sectional and prospective cohort study of the role of schools in the SARS-CoV-2 second wave in Italy. *Lancet Reg. Health Eur.* **5**, 100092 (2021).
- Goldhaber, D. et al. To what extent does in-person schooling contribute to the spread of COVID-19? Evidence from Michigan and Washington. Preprint at <http://www.nber.org/papers/w28455> (2021).
- Harris, D. N., Ziedan, E. & Hassig, S. The effects of school reopenings on COVID-19 hospitalizations. Preprint at <https://www.reachcentered.org/publications/the-effects-of-school-reopenings-on-covid-19-hospitalizations> (2021).
- Hsiang, S. et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).
- Ispphording, I. E., Lipfert, M. & Pestel, N. Does re-opening schools contribute to the spread of SARS-CoV-2? Evidence from staggered summer breaks in Germany. *J. Public Econ.* **198**, 104426 (2021).
- Iwata, K., Doi, A. & Miyakoshi, C. Was school closure effective in mitigating coronavirus disease 2019? Time-series analysis using Bayesian inference. *Int. J. Infect. Dis.* **99**, 57–61 (2020).
- Lin, P. Z. & Meissner, C. M. Health vs. wealth? Public health policies and the economy during COVID-19. *Covid Econ.* **14**, 85–106 (2020).
- Plümper, T. & Neumayer, E. Summer school holidays and the growth rate in SARS-CoV-2 infections across German districts. Preprint at <https://www.medrxiv.org/content/10.1101/2020.10.11.2010773v2> (2020).
- von Bismarck-Osten, C., Borusyak, K. & Schönberg, U. The role of schools in transmission of the SARS-CoV-2 virus: quasi-experimental evidence from Germany. Preprint at https://www.cream-migration.org/publ_uploads/CDP_22_20.pdf (2020).

49. Yoon, Y., Kim, K.-R., Park, H., Kim, S. & Kim, Y.-J. Stepwise school opening and an impact on the epidemiology of COVID-19 in the children. *J. Korean Med. Sci.* **35**, e414 (2020).
50. Imbens, G. W. & Rubin, D. B. *Causal Inference for Statistics, Social, and Biomedical Sciences* (Cambridge University Press, 2015).
51. Angrist, J. D. & Pischke, J.-S. *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton Univ. Press, 2009).
52. Ho, D. E., Imai, K., King, G. & Stuart, E. A. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit. Anal.* **15**, 199–236 (2007).
53. Forastiere, L., Airola, E. M. & Mealli, F. Identification and estimation of treatment and interference effects in observational studies on networks. *J. Am. Stat. Assoc.* **116**, 901–918 (2021).
54. Sun, K. et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371**, eabe2424 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Methods

Data. *Treatment variables.* We focused on schools established by municipalities. Usually, each school is established by one municipality. However, sometimes a few municipalities jointly establish a ‘school authority’ and operate the school together. In such a case, we regarded the school as belonging to each of the municipalities. The sources that we used from the MEXT treat the ‘founder’ of schools as the unit of observation, that is, a municipality and a school authority.

Most of the schools we studied are elementary schools (K1–6) and junior high schools (K7–9). We excluded high schools (K10–12, most of which are established by prefectures), universities (most of which are established by the national government and private entities) and nurseries.

The MEXT conducted surveys on school closures across municipalities nine times in 2020. (Besides the websites we refer to below, see the MEXT, ‘shingata korona uirusu ni kanrenshita kansensho taisaku ni kansuru tao ni tsuite: shinshaku joho (on the measures to address COVID-19: news); https://www.mext.go.jp/a_menu/coronavirus/index_00012.html; accessed on 17 June 2021). We collapsed two of them into the 16 March survey as detailed below. If a survey does not report whether all or some of schools are closed in a municipality (for example, the school authority that the municipality is part of did not respond to the survey), we coded the treatment variable as having a missing value in the municipality.

4 March. The source is the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no sho-chu-koto gakkoto ni okeru rinji kyugyo no jokyo ni tsuite (on the situation of temporary closures of elementary, junior high and high schools to address COVID-19)’, 4 March 2020 (https://www.mext.go.jp/content/20200304-mxt_kouhou02-000004520_1.pdf; accessed on 7 August 2020). The survey lists municipalities that opened all of their elementary and/or junior high schools. We supposed that other municipalities and school authorities closed all of their elementary and junior high schools.

16 March. The sources are the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no shogakko chugakko kotogakko oyobi tokubetsushien gakkoto ni okeru rinji kyugyo no jokyo ni tsuite (on the situation of temporary closures of elementary, junior high, high and special education schools to address COVID-19)’, 16 March 2020 (https://www.mext.go.jp/content/20200316_mxt_kouhou02_000004520-1.pdf; survey 1, accessed on 7 August 2020) and the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no shogakko no rinji kyugyo ni kanrenshita kodomo no ibasho no kakuhoto ni kansuru kaku jichitai no torikumi jokyo ni tsuite (on the situation of each municipality’s measure to prepare for the children’s place concerned with temporary closures of elementary and other schools to address COVID-19)’, 17 March 2020 (https://www.mext.go.jp/content/20200317-mxt_kouhou02-000004520_1.pdf; survey 2, accessed on 7 August 2020). Survey 1 lists the founders that planned to open all of their elementary and/or junior high schools as of 16 March. Survey 2 lists the founders that were not listed in survey 1 and ended closures of their elementary and/or junior high schools by 15 March. In sum, we can say that both groups of founders opened all of their schools as of 16 March. We supposed that other founders closed all of their schools as of 16 March. Hamamatsu City has a missing value because surveys 1 and 2 contradict each other.

6 April. The source is the MEXT, ‘shingata korona uirusu kansensho taisaku ni kansuru gakko no shingakki kaishi jokyoto ni tsuite (on the situation of the beginning of schools’ new term concerned with measures to address COVID-19)’, 6 April 2020 (https://www.mext.go.jp/content/20200407-mxt_kouhou01-000006421_1.pdf; accessed on 29 June 2020). The survey lists the founders that closed their schools. We supposed that other founders opened all of their schools.

10 April. The source is the MEXT, ‘shingata korona uirusu kansensho taisaku ni kansuru gakko no shingakki kaishi jokyoto ni tsuite (on the situation of the beginning of schools’ new term concerned with measures to address COVID-19)’, 10 April 2020 (https://www.mext.go.jp/content/20200413-mxt_kouhou01-000006421_1.pdf; accessed on 29 June 2020). The survey lists founders that closed their schools. We supposed that other founders opened all of their schools.

16 April. The source is the MEXT, ‘gakushu shidoto torikumi jokyo chosa (survey on the status of learning guidance)’, 16 April 2020. We obtained the source data files from the MEXT on 24 December 2020 with their permission. The aggregated results are reported at the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no gakko no rinji kyugyo ni kanrenshita koritsu gakkō niokeru gakushu shidoto no torikumi jokyo ni tsuite (on the status of learning guidance in public schools concerned with temporary closures of schools to address COVID-19)’, 16 April 2020 (https://www.mext.go.jp/content/20200421-mxt_kouhou01-000004520_4.pdf; accessed on 29 October 2020). The MEXT asked founders to respond if they closed their schools. Thus, we supposed that founders who did not respond opened all of their schools.

22 April. The source is the MEXT, ‘rinji kyugyo jisshi jokyo chosa (survey on the status of temporary school closures)’, 22 April 2020. We obtained the source data files from the MEXT on 24 December 2020 with their permission. The aggregated

results are reported at the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no gakko niokeru rinji kyugyo no jisshi jokyo ni tsuite (on the status of temporary closures of schools to address COVID-19)’, April 22, 2020 (https://www.mext.go.jp/content/20200424-mxt_kouhou01-000006590_1.pdf; accessed on 14 October 2020). As for surveys after 22 April, the source data files that the MEXT provides clarify the information not only about elementary and junior high schools but also about ‘obligatory education schools’ (K1–9), ‘secondary education schools’ (K7–12) and ‘special education schools’ (K1–12 and preschool for those who have disabilities and need special help), all of which we take into account in making the treatment variables. In principle, all founders are supposed to respond, although in practice, some failed to do so and their treatment variable (or the treatment variable of municipalities whose school authorities failed to respond) has a missing value due to data unavailability.

11 May. The source is the MEXT, ‘rinji kyugyo jisshi jokyo chosa (survey on the status of temporary school closures)’, 11 May 2020. We obtained the source data file from the MEXT on 24 December 2020 with their permission. The aggregated results are reported at the MEXT, ‘shingata korona uirusu kansensho taisaku no tame no gakko ni okeru rinji kyugyo no jisshi jokyo ni tsuite (on the status of temporary closures of schools to address COVID-19)’, 11 May 2020 (https://www.mext.go.jp/content/20200513-mxt_kouhou02-000006590_2.pdf; accessed on 14 October 2020). One municipality has a missing value because its answer is inconsistent.

1 June. The source is the MEXT, ‘gakko saikai jokyo chosa (survey on the status of school reopenings)’, 4 June 2020. We obtained the source data file from the MEXT on 24 December 2020 with their permission. The aggregated results are reported at the MEXT, ‘shingata korona uirusu kansensho ni kansuru gakko no saikai jokyo ni tsuite (on the status of reopening of schools concerned with COVID-19)’, 1 June 2020 (https://www.mext.go.jp/content/20200603-mxt_kouhou01-000004520_4.pdf; accessed on 14 October 2020).

Outcome variables. Our outcome variables were the daily number of newly confirmed cases per 100,000 residents. We collected the daily number of newly confirmed cases from prefectures’ websites, whose URLs are displayed in Extended Data Fig. 1 and Supplementary Table 1 (accessed on 12 August 2020, although we updated the URLs (but not the data) on 22 March 2021; an exception is Tokyo Prefecture, whose URL was accessed on 15 January 2021). The values of the outcome variables were only available for all municipalities in 26 of 47 prefectures. In addition, we coincidentally obtained municipal-level data from Tokushima Prefecture through a Freedom of Information Act (FOIA) request. Therefore, we only analyzed 27 ‘target’ prefectures.

We were careful when analyzing data from Tokyo Prefecture. The prefecture only started reporting the cumulative number of confirmed cases in each of its 62 municipalities after 31 March. By taking the difference between two consecutive days, we were able to obtain the daily number of newly confirmed cases, but only after 1 April. It is worth noting that in 17 observations, these differences are equal to −1, perhaps because Tokyo corrected the data from the day before; we changed these values to zero. Accordingly, we were able to calculate these values for the past 7 d only after 8 April. Therefore, Tokyo prefecture was included only for the treatments as of 10 April and after. This is why the number of target prefectures is 26 for 6 April and earlier treatments and 27 for 10 April and later treatments.

Data on population (including foreigners) is as of 1 January 2020 (latest available data). The source is ‘jumin kihon daicho ni motoduku jinko, jinko dotai oyobi setaisu (population, its movement and the number of households based on residential basic book)’, 1 January 2020 (https://www.soumu.go.jp/main_sosiki/jichi_gyousei/daityo/jinkou_jinkoudoutai-setaisuu.html; accessed on 27 January 2021).

Because the daily number of newly confirmed cases in a municipality was small and our study period (26 February to 1 June) was at an early stage of the pandemic, the number of deaths were much smaller and thus not analyzed. For the same reason, we cannot calculate the reproduction number.

Covariates. Besides the past treatment variables and the set of prefecture dummies, we used 49 covariates. They included: the sum of all the outcome variables before the survey date; the outcome variables from the past 7 d; the number of municipalities covered by the public health center that is in charge of a municipality; five demographic variables (population, population density, the young, the old and densely inhabited districts); seven commuting variables (in-migrants, out-migrants, commuters from other municipalities in the same prefecture, commuters from other prefectures, commuters to other municipalities in the same prefecture, commuters to other prefectures and daytime population); four geographic variables (inhabitable area size, the number of bordering municipalities, latitude and longitude); income; four variables on the municipal government’s fiscal situation (financial solidity index, total revenue, local tax and non-transferred revenue); four education variables (elementary school pupils, elementary school pupils per school, junior high school students and junior high school students per school); four labor variables (labor force, unemployment, primary industry employment and secondary industry employment); five medical

variables (hospitals, medical clinics, beds of hospitals, beds of medical clinics and physicians); three climatic variables (precipitation, daylight hours and average temperature); and three mayoral variables (age, number of terms and days since last election). In the Supplementary Information (main analysis), we elaborate on why we chose these covariates. We divided the description of the covariates by their sources.

Government statistics. Extended Data Fig. 2 summarizes how to calculate most of the covariates. The numerators and denominators are original variables retrieved from the source, National Statistics Center, ‘tokei de miru todofukken shichoson no sugata (shakai jinko tokei taikei) (system of social and demographic statistics (municipality data))’ (<https://www.e-stat.go.jp/regional-statistics/ssdview/municipality>; accessed on 5 and 15 October 2020 and 9 February 2021). For the financial solidity index (D2201) of the 23 special wards in Tokyo Prefecture, we referred to Tokyo Metropolitan Government, Bureau of General Affairs, Local Administration Division, ‘tokubetsu futsu kaikei kessan no jykyo (the state of the settled general accounts of special wards), 2017’ (<https://www.soumu.metro.tokyo.lg.jp/05gyousei/gyouzaisei/new/30tuita/29soukatsu-ku-shi/29gaiyo.xls>; accessed on 8 February 2021). Below, the variables are explained and referred to by their official codes in the dataset (Supplementary Information; data, covariates, government statistics). The year for each variable is in parentheses. We refer to the latest available year. If the reference year is 2015, the original variable is in the census. If the unit is not the number of people or items, we mention it in the square brackets at the end of the line.

A1101 Population (2015)
A1301 Young population (14 and younger) (2015)
A1303 Old population (65 and older) (2015)
A1801 Densely inhabited districts population (2015)
A2301 Population (2016–2018)
A5103 In-migrants (2018)
A5104 Out-migrants (2018)
A6103 Workers and students to other municipalities in the same prefecture (2015)
A6104 Workers and students to other prefectures (2015)
A6105 Workers and students from other municipalities in the same prefecture (2015)

A6106 Workers and students from other prefectures (2015)
A6108 Daytime population relative to the population (percentage) (2015)
B1101 Total land area (excluding the northern territories and Takeshima Island) (2018) [ha]

B1103 Inhabitable area (2018) [ha]
C120110 Taxable income (2018) [1,000 yen]
C120120 Number of tax debtors (per income) (2018)
D2201 Financial solidity index (2017)
D3201 Total revenue (2017) [1,000 yen]
D320101 Local tax revenue (2017) [1,000 yen]
D3202 Non-transferred revenue (2017) [1,000 yen]
E2101 Number of elementary schools (2018)
E2501 Number of elementary school students (2018)
E3101 Number of junior high schools (2018)
E3501 Number of junior high school students (2018)
F1101 Population in labor force (2015)
F1102 Number of employed persons (2015)
F1107 Number of unemployed persons (2015)
F2201 Number of persons employed in primary industry (2015)
F2211 Number of persons employed in secondary industry (2015)
I5101 Number of hospitals (2017)
I5102 Number of medical clinics (2017)
I5211 Beds of hospitals (2017)
I5212 Beds of medical clinics (2017)
I6100 Number of physicians (2016)

We matched on neither census population in 2015 (the variable A1101) nor population in 2016–2018 (the variable A2301) but the latest population (including foreigners) as of 1 January 2020, which we introduced in the previous subsection. As for the variable A1801, we set missing values of this variable to zero, supposing that such municipalities have no densely inhabited districts. Other variables have no missing values.

In 13 municipalities in Fukushima Prefecture (Futaba, Hirono, Iitate, Iwaki, Katsurao, Kawamata, Kawauchi, Minami-soma, Namie, Naraha, Okuma, Tamura and Tomioka), which are eligible for the Nuclear Power Accident Evacuation Special Act, residents and students do not necessarily live in the municipalities. Therefore, we excluded these municipalities from our analysis.

Public health centers. Another covariate is the log of the number of municipalities that the public health center in charge of a given municipality is in charge of. The source is the Ministry of Health, Labor and Welfare, ‘todofukken betsushi shikuchoson fugo oyobi hokenjo fugo (municipality codes and public health center codes by prefecture)’ (https://www.data.go.jp/data/dataset/mhlw_20170316_0002; accessed on 29 January 2021). We refer to this source when we use health center fixed effects

as well. Although the ‘Matsue Public Health Center Established by Matsue City and Shimane Prefecture’ is assigned different public health center codes in Matsue and Yasugi Cities, which the center is in charge of, we supposed both cities had the same public health center. There are no missing values for this variable.

Meteorological data. We also matched on the annual normal values (average between 1981–2010) of precipitation (mm), daylight hours and average temperature (degrees Celsius). The source is ref.⁵⁵, which contains values of the three variables by 1-km-square mesh. We take their averages across 1-km-square grid cells in each municipality. A grid cell at the border of municipalities was assigned to the municipality that included the center of the grid cell. Ogasawara Village in Tokyo Prefecture, a far remote island, had missing values for these three meteorological variables and thus was excluded from our analysis.

Geographical data. We downloaded the shape files of all municipalities from the Ministry of Land, Infrastructure, Transport and Tourism, National Spatial Planning and Regional Policy Bureau, National Land Information Division, ‘kokudo suchi joho daunrodo (downloadable files of land numerical information)’ as of 1 January 2020 (<https://nlftp.mlit.go.jp/ksj/index.html>; accessed on 27 January 2021). Based on the files, we obtained latitude, longitude and the number of neighbor municipalities. We took the logarithm of the number of neighbor municipalities after adding one to it. Some municipalities (usually, islands) had no neighbor municipality. There were no missing values of these variables.

Mayoral data. We used data from Ichini, ‘senkyo dotto komu (election dot-com)’ (<https://go2senkyo.com/>; accessed on 28 April 2021) for information on mayors including their ages, number of terms and days since their last election. We measured them as of the survey date for each treatment variable. We filled in the few missing values of variables using newspaper coverage of elections.

Matching. We implement genetic matching⁵⁶ with replacement by using the MatchIt package⁵⁷ in the statistical computing environment R⁵⁸, which calls functions from the Matching package⁵⁹. We matched on the covariates discussed in the previous section. Upon advice by ref.⁶⁰, as a metric of the distance between municipalities, we used the Mahalanobis distance. Following ref.⁵⁹, we set the arguments pop.size and nboots at 1,000. We kept other options at their default values.

Statistical analyses: main analysis. To derive the confidence intervals of the ATC values, we utilized cluster-robust standard errors by using the lmtest and sandwich packages where clusters were a pair of matched municipalities and a municipality^{61–63}.

Statistical analyses: robustness checks. Negative binomial regression. We implemented negative binomial regression using the MASS package⁶⁴.

Public health center fixed effects. Depending on the survey date, there were only 25 to 139 municipalities covered by public health centers that had responsibility for both treated and control municipalities, which contributed to the identification of the treatment effects (Supplementary Table 5). Therefore, we did not regress on any covariate. We implemented fixed-effects models using the lfe package⁶⁵. Standard errors were clustered by public health center.

Inverse probability weighting. We estimated the propensity score using the covariate balancing propensity score algorithm⁶⁶, which is implemented by the CBPS package. We estimated the ATC values by way of the WeightIt package^{51,67}. When calculating standard errors, we considered weight by way of the survey package⁶⁸.

Conditioning on neighbors. If any of the neighbor municipalities had a missing value of the treatment variable or if a municipality such as an island had no neighbor municipalities, such municipalities were excluded from the analysis. We considered neighbors outside the 27 target prefectures as well.

A smaller set of covariates. In addition to the past treatment variables and the set of prefecture dummies, we still matched on the following covariates: the total number of COVID-19 cases before the survey date; the number of COVID-19 cases in each of the past 7 d; the number of municipalities covered by the public health center that is in charge of a municipality; population; population density; the old; densely inhabited population districts; daytime population; income; financial solidity index; local tax; elementary school pupils; elementary school pupils per school; junior high school students; junior high school students per school; hospitals; medical clinics; and physicians.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data except the four surveys (as of 16 and 22 April, 11 May and 1 June 2020) from the MEXT are deposited to Harvard Dataverse as K.F., C.T.M. and K.N., 2021, ‘replication data for: no causal effect of school closures in Japan on the spread of

COVID-19 in spring 2020,⁷ at <https://doi.org/10.7910/DVN/N803UQ>. All the data sources are detailed in the Methods. Information regarding the undisclosed four surveys is as follows:

Reasons for controlled access: the MEXT does not allow users of the data to disclose it.

Precise conditions of access (including contact details for access requests): it is necessary to get permission of the MEXT (Monbusho, shoto chuto kyoiku kyoku, kenkyo kyoiku shokuhaku ka (Ministry of education, culture, sports, science and technology, Bureau of elementary and secondary education, Health and dietary education division).

A time frame for response to requests: we submitted our FOIA request to the MEXT on 30 October 2020 and obtained the source data files from the MEXT on 24 December 2020 with their permission.

Details of any restrictions imposed on data use via data use agreements: users should not disclose the data and cannot report analysis of the data so that readers can know the school closure status of a municipality.

Code availability

The replication code is deposited to Harvard Dataverse at <https://doi.org/10.7910/DVN/N803UQ>.

References

55. Kishio Gyomu Shien Senta. *Messhu Heinench 2010 (Mesh Normal Value 2010)* (Kishio Gyomu Shien Senta, 2016).
56. Diamond, A. & Sekhon, J. S. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Rev. Econ. Stat.* **95**, 932–945 (2013).
57. Ho, D., Imai, K., King, G. & Stuart, E. A. Matchit: nonparametric preprocessing for parametric causal inference. *J. Stat. Softw.* **42**, 1–28 (2011).
58. R Development Core Team. R: a language and environment for statistical computing (version 4.0.3) (2020).
59. Sekhon, J. S. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *J. Stat. Softw.* **42**, 1–52 (2011).
60. King, G. & Nielsen, R. Why propensity scores should not be used for matching. *Polit. Anal.* **27**, 435–454 (2019).
61. Zeileis, A. & Hothorn, T. Diagnostic checking in regression relationships. *R News* **2**, 7–10 (2002).
62. Zeileis, A. Econometric computing with HC and HAC covariance matrix estimators. *J. Stat. Softw.* **11**, 1–17 (2004).
63. Zeileis, A., Köll, S. & Graham, N. Various versatile variances: an object-oriented implementation of clustered covariances in R. *J. Stat. Softw.* **95**, 1–36 (2020).
64. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. 4th Ed. (Springer, 2002).
65. Gauvreau, S. Ife: linear group fixed effects. R package version 2.8-3 <https://CRAN.R-project.org/package=ife> (2019).
66. Imai, K. & Ratkovic, M. Covariate balancing propensity score. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **76**, 243–263 (2014).
67. Greifer, N. WeightIt: weighting for covariate balance in observational studies. R package version 0.11.0 <https://CRAN.R-project.org/package=WeightIt> (2021).
68. Lumley, T. survey: analysis of complex survey samples. R package version 4.0 <http://r-survey.r-forge.r-project.org/survey/> (2020).

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science KAKENHI (grant no. JP19K21683). We obtained part of the survey data from the MEXT. The funder and the MEXT had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

K.F. conceived the study. K.F., C.T.M. and K.N. collected the data. C.T.M. analyzed the data. K.F. wrote the first draft of the manuscript. All authors interpreted the results, contributed to writing and approved the final version for submission.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41591-021-01571-8>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-021-01571-8>.

Correspondence and requests for materials should be addressed to Kentaro Fukumoto.

Peer review information *Nature Medicine* thanks Carlos del Rio and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Alison Farrell and Jennifer Sargent are the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

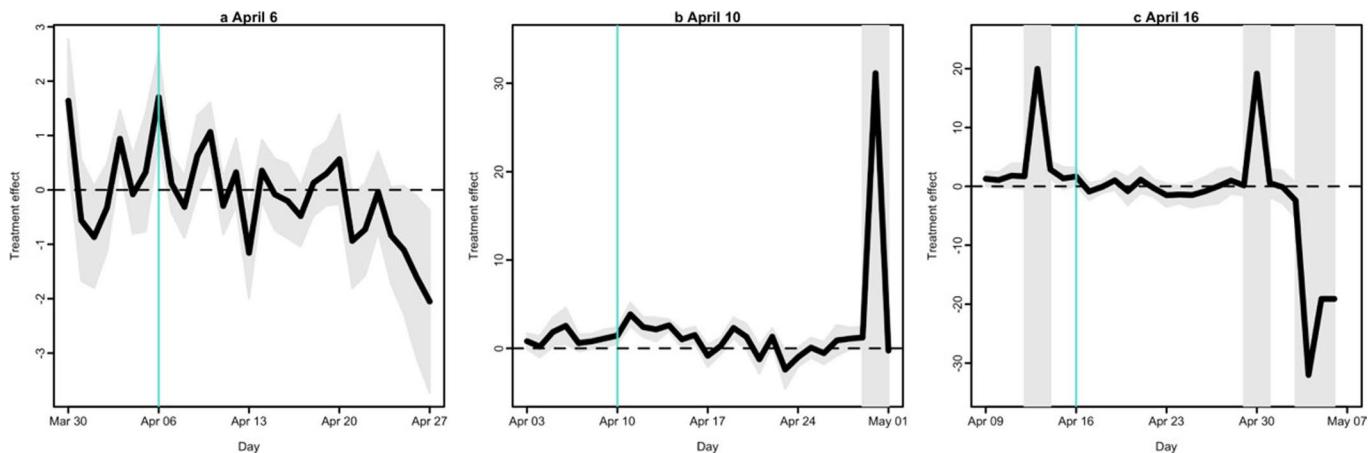
Reprints and permissions information is available at www.nature.com/reprints.

Prefecture	Source
Iwate	https://www.pref.iwate.jp/kurashikankyou/iryou/covid19/1029635/index.html
Yamagata	https://www.pref.yamagata.jp/090016/bosai/kochibou/kikikanri/covid19/shingata_corona2.html
Fukushima	https://www.pref.fukushima.lg.jp/sec/21045c/fukushima-hasseijyoukyou.html
Tochigi	http://www.pref.tochigi.lg.jp/e04/welfare/hoken-eisei/kansen/hp/coronakensahasseijyoukyou.html
Saitama	https://www.pref.saitama.lg.jp/a0701/covid19/jokyo.html
Chiba	https://www.pref.chiba.lg.jp/shippei/press/2019/ncov-index.html
Tokyo	https://www.bousai.metro.tokyo.lg.jp/taisaku/saigai/1010035/index.html
Niigata	https://www.pref.niigata.lg.jp/site/shingata-corona/256362836.html
Toyama	http://www.pref.toyama.jp/cms_sec/1205/kj00021798.html
Fukui	https://covid19-fukui.com/
Gifu	https://www.pref.gifu.lg.jp/site/covid19/26547.html
Aichi	https://www.pref.aichi.jp/site/covid19-aichi/
Shiga	https://stopcovid19.pref.shiga.jp/
Osaka	https://covid19-osaka.info/
Tottori	https://www.pref.tottori.lg.jp/291425.htm
Shimane	https://www.pref.shimane.lg.jp/bousai_info/bousai/kikikanri/shingata_taisaku/new_coronavirus_portal.html
Okayama	https://www.pref.okayama.jp/page/667843.html
Hiroshima	https://www.pref.hiroshima.lg.jp/site/hcdc/covid19-kanja.html
Yamaguchi	https://yamaguchi.stopcovid19.jp/
Tokushima	https://www.pref.tokushima.lg.jp/ippannokata/kenko/kansensho/5034012#20
Kagawa	https://www.pref.kagawa.lg.jp/yakumukansen/kansenshoujouhou/kansen/se9si9200517102553.html
Ehime	https://ehime-covid19.com/
Saga	https://www.pref.saga.lg.jp/kiji00373220/index.html
Nagasaki	https://www.pref.nagasaki.jp/bunrui/hukushi-hoken/kansensho/corona_nagasaki/corona_nagasaki_shousai/
Oita	https://www.pref.oita.jp/site/covid19-oita/covid19-pcr.html
Miyazaki	https://www.pref.miyazaki.lg.jp/fukushihoken/covid-19/kenmin/20200804143434.html
Kagoshima	https://www.pref.kagoshima.jp/ae06/kenko-fukushi/kenko-iryo/kansen/kansensho/coronavirus.html

Extended Data Fig. 1 | Sources of the daily number of newly confirmed cases: analyzed prefectures. The daily number of newly confirmed cases was collected from these prefectures' websites on 12 August 2020. The URLs (but not the data) were updated on 22 March 2021.

Variable name	log?	Numerator	Denominator
Young Population (<15)		A1301	÷ A1101
Old Population (=> 65)		A1303	÷ A1101
Densely inhabited districts population		A1801	÷ A1101
Population Density	log(A2301 (2018)	÷ B1101)
In-migrants		A5103	÷ A2301 (2018)
Out-migrants		A5104	÷ A2301 (2018)
Commuters from other municipalities in the same prefecture		A6103	÷ A1101
Commuters from other prefectures		A6104	÷ A1101
Commuters to other municipalities in the same prefecture	log(A6105 +1	÷ A1101)
Commuters to other prefectures	log(A6106 +1	÷ A1101)
Daytime Population	log(A6108	÷ 100)
Livable area		B1101	÷ B1103
Income	log(C120120	÷ C120110)
Financial solidity index	log(D2201)
Total revenue	log(D3201	÷ A2301 (2017))
Local tax		D320101	÷ D3201
Non transferred revenue		D3202	÷ D3201
Elementary school pupils		E2501	÷ A2301 (2018)
Elementary school pupils per school*	log(E2501 +1	÷ E2101)
Junior high school students		E3501	÷ A2301 (2018)
Junior high school students per school*	log(E3501 +1	÷ E3101)
Labor Force		F1101	÷ A1101
Unemployment		F1107	÷ F1101
1st Industry Employment		F2201	÷ F1102
2nd Industry Employment		F2211	÷ F1102
Hospitals		I5101	÷ A2301 (2017)
Medical clinics		I5102	÷ A2301 (2017)
Beds of hospitals		I5211	÷ A2301 (2017)
Beds of medical clinics (2017)		I5212	÷ A2301 (2017)
Physicians		I6100	÷ A2301 (2016)

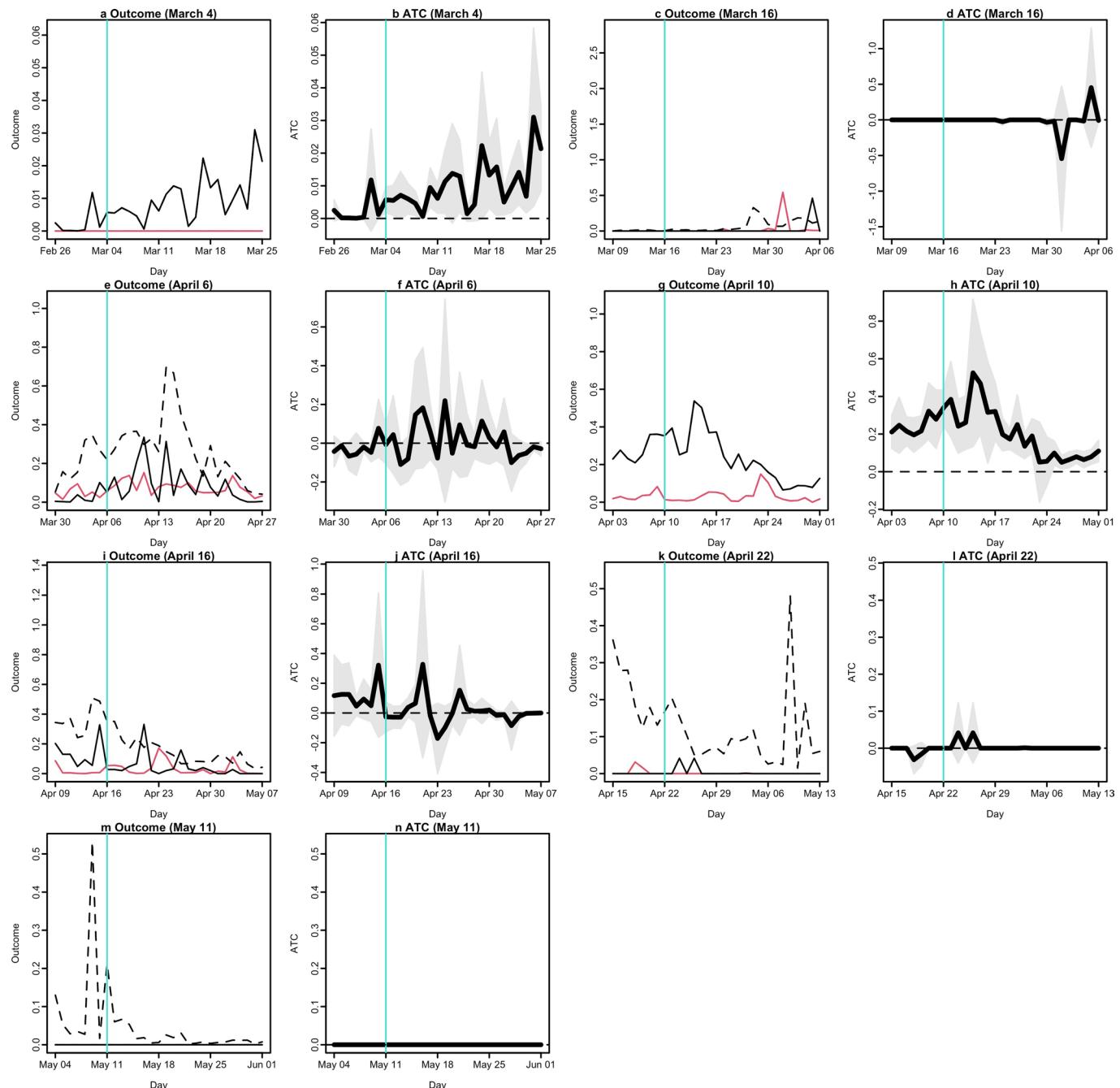
Extended Data Fig. 2 | Definition of covariates. Numerators and denominators refer to the variable numbers in National Statistics Center, 'Tokei de miru todofukan shichoson no sugata (shakai jinko tokei taikei) [System of social and demographic statistics (municipality data)]' <https://www.e-stat.go.jp/regional-statistics/ssdsview/municipality> (accessed on October 5 and 15, 2020 and February 9, 2021). The year in parenthesis next to the variable A2301 refers to as of when the variable refers to the population. The value of the two variables with asterisk * is equal to zero when the denominator takes the value of zero.



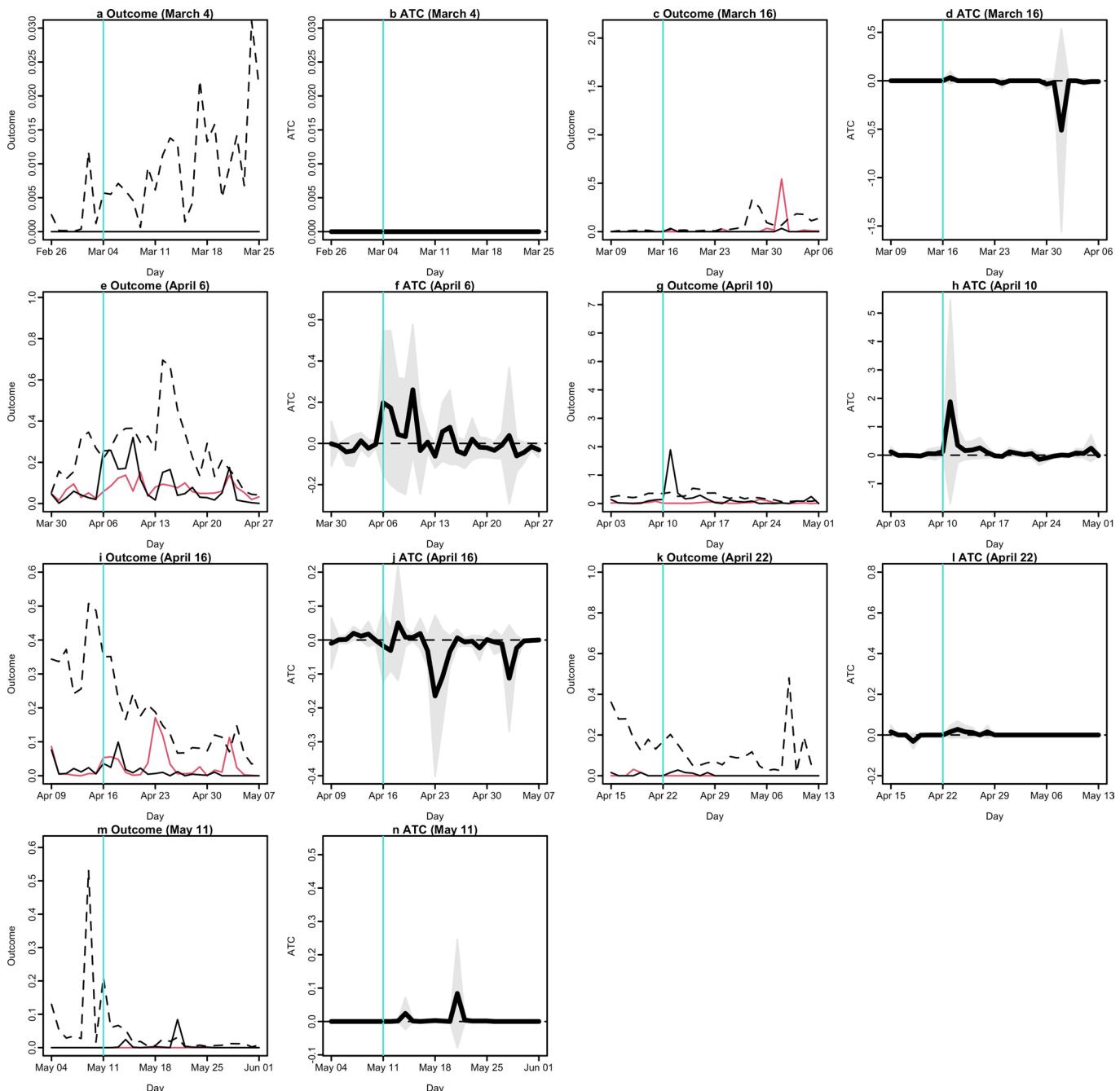
d Irregular results of treatment effects by negative binomial regression

Treatment	Outcome	Point estimate	Standard error	Confidence interval
April 10	April 30	31.1	620,041.4	-1,215,227.6 1,215,289.9
April 16	April 16	20.0	4,622.0	-9,039.0 9,078.9
	April 30	19.1	4,622.0	-9,039.8 9,078.1
	May 4	-32.0	3,697,991.9	-7,247,963.0 7,247,899.0
	May 5	-19.1	8,005.0	-15,708.6 15,670.4
	May 6	-19.1	8,005.0	-15,708.6 15,670.4
	May 7	(the estimation procedure did not converge)		

Extended Data Fig. 3 | Treatment effects: negative binomial regression. In each panel, the horizontal axis indicates dates in 2020, and the vertical axis represents the coefficient estimates of the treatment variable. The vertical turquoise line corresponds to the survey date. The thick black line indicates the point estimates of the treatment variable's coefficients, and the shaded grey area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters are a pair of matched municipalities and a municipality. **a**, April 6. **b**, April 10. **c**, April 16. **d**, Irregular results of treatment effects by negative binomial regression. For other survey dates, we cannot estimate negative binomial regression parameters.



Extended Data Fig. 4 | Average outcomes and ATCs: inverse probability weighting. In each panel, the horizontal axis indicates dates in 2020, and the vertical turquoise line corresponds to the survey date. In the outcome panels (**a, c, e, g, i, k, m**), the vertical axis represents the average number of confirmed cases per 100,000 residents, the black and red lines correspond to the weighted average outcomes of the treated and control municipalities, respectively, and the dotted black line represents the un-weighted average outcomes of treated municipalities. In the ATC panels (**b, d, f, h, j, l, n**), the vertical axis represents ATCs, the thick black line indicates the point estimates of ATCs, and the shaded grey area presents the 95% confidence intervals. We calculate standard errors, taking into consideration weight. **a**, outcome, March 4. The dotted black line overlaps the solid black line and is thus invisible. **b**, ATC, March 4. **c**, outcome, March 16. **d**, ATC, March 16. **e**, outcome, April 6. **f**, ATC, April 6. **g**, outcome, April 10. The dotted black line overlaps the solid black line and is thus invisible. **h**, ATC, April 10. **i**, outcome, April 16. **j**, ATC, April 16. **k**, outcome, April 22. **l**, ATC, April 22. **m**, outcome, May 11. **n**, ATC, May 11.



Extended Data Fig. 5 | Average outcomes and ATCs: a smaller set of covariates. In each panel, the horizontal axis indicates dates in 2020, and the vertical turquoise line corresponds to the survey date. In the outcome panels (**a, c, e, g, i, k, m**), the vertical axis represents the average number of confirmed cases per 100,000 residents, the black and red lines correspond to the average outcomes of the matched treated and control municipalities, respectively, and the dotted black line represents the average outcomes of all treated municipalities, both matched and unmatched. In the ATC panels (**b, d, f, h, j, l, n**), the vertical axis represents ATCs, the thick black line indicates the point estimates of ATCs, and the shaded grey area presents the 95% confidence intervals. Cluster-robust standard errors are utilized where clusters are a pair of matched municipalities and a municipality. **a**, outcome, March 4. **b**, ATC, March 4. **c**, outcome, March 16. **d**, ATC, March 16. **e**, outcome, April 6. **f**, ATC, April 6. **g**, outcome, April 10. **h**, ATC, April 10. **i**, outcome, April 16. **j**, ATC, April 16. **k**, outcome, April 22. **l**, ATC, April 22. **m**, outcome, May 11. **n**, ATC, May 11.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection No software was used for data collection purposes.

Data analysis R: A Language and Environment for Statistical Computing (Version 4.0.3). The R packages we use are tidyverse (version 1.3.1), checkpoint (version 0.4.10), MatchIt (version 4.1.0), Matching (version 4.9-9), rgenoud (version 5.8-3.0), lmtest (version 0.9-38), sandwich (version 3.0-1), lfe (version 2.8-3), MASS (version 7.3-53), CBPS (version 0.22), WeightIt (version 0.11.0), survey (version 4.0), and cobalt (version 4.3.1). The replication code is deposited to Harvard Dataverse, <https://doi.org/10.7910/DVN/N803UQ>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All data except the four surveys (as of April 16, 22, May 11, and June 1, 2020) from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) is deposited to Harvard Dataverse as Fukumoto, Kentaro; Charles T. McClean; Kuninori Nakagawa, 2021, "Replication Data for: No causal effect of school closures in Japan on the spread of COVID-19 in spring 2020," at <https://doi.org/10.7910/DVN/N803UQ>. All the data sources are detailed in Methods section, Data subsection. Information of the non-disclosed four surveys is as follows.

Reasons for controlled access: MEXT does not allow users of the data to disclose it.

Precise conditions of access (including contact details for access requests): it is necessary to get permission of MEXT (Monbusho, shoto chuto kyoiku kyoku, kenkyo kyoiku shokuhaku ka [Ministry of education, culture, sports, science and technology, Bureau of Elementary and secondary education bureau, Health and dietary education division] (address: 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo-to 100-8959 Japan).

A timeframe for response to requests: We submitted our FOIA request to MEXT on October 30, 2020 and obtained the source data files from MEXT on December 24, 2020 with their permission.

Details of any restrictions imposed on data use via data use agreements: Users should not disclose the data and cannot report analysis of the data so that readers can know the school closure status of a municipality.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

For every set of survey date and outcome date, we implement a cross-section analysis across 847 municipalities where we perform matching and regress the outcome variable on the treatment variable using only the matched municipalities. The survey dates are March 4, 16, April 6, 10, 16, 22, and May 11. The outcome dates range from February 26 to June 1, 2020.

Research sample

847 municipalities in Japan. They are not necessarily representative. They are chosen because the number of COVID-19 cases is available.

Sampling strategy

847 municipalities in Japan are chosen because data on the daily number of COVID-19 cases is available.

Data collection

We download the data from various websites. Those who maintain the data recorded the data.

Timing

From February 26 to June 1, 2020.

Data exclusions

No data were excluded from the analyses.

Non-participation

This item is not applicable because our study is observational.

Randomization

Municipalities were not allocated into experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging