WINPEC Working Paper Series No.E2107
May 2021

Household Expenditures and the Effective Reproduction Number in Japan: Regression Analysis

Hajime Tomura

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

Household Expenditures and the Effective Reproduction Number in Japan: Regression Analysis

Hajime Tomura*

May 8, 2021

Abstract

Daily estimates of the effective reproduction number for new coronavirus based on reporting dates are regressed on real household expenditures per household on eating out, traveling, and apparel shopping, as well as mobility in public transportation, using publicly available daily nationwide data from February 15, 2020, to February 1, 2021 in Japan. The effects of absolute humidity, the declaration of states of emergency, and the year-end and new-year holiday period are controlled through dummy variables. The lagged infectious effect of economic activities due to incubation periods is also taken into account. Estimated regression coefficients indicate that real household expenditures for cafe and bar had larger effects on the effective reproduction number per value of spending than the other types of household expenditures in explanatory variables during the sample period. Thus, a loss of aggregate demand is minimized if the effective reproduction number is lowered by restricting only household consumption of cafe and bar. The posterior means of simulations based on the estimated regression coefficients, however, imply that even if a self-restraint on packaged domestic travels and an endogenous decline in mobility are taken into account, it will be necessary to cut household consumption of cafe and bar by more than 80% of the 2019 level, in order to keep the effective reproduction number below one on average.
JEL codes: E21, I18.
Keywords: new coronavirus; effective reproduction number; consumption; mobility.

*Faculty of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjyuku-ku, Tokyo, 169-8050, Japan. Email: htomura@waseda.jp. I thank Hiroshi Fujiki, Munechika Katayama, Keisuke Otsu, and seminar participants at Waseda University for their valuable comments.

1 Introduction

Japanese pharmaceutical companies are still in the process of developing new-coronavirus vaccines as of April 2021. Thus, the country will have to wait for the arrival of vaccines, including imports, and contain the spread of new-coronavirus infection by intervening in economic activities until then. To figure out a cost-effective intervention in terms of a loss of aggregate demand, I regress the effective reproduction number on real household expenditures per household for eating out, traveling, and apparel shopping, as well as a measure of mobility in public transportation, using publicly available daily nationwide data in Japan. These real household expenditures are included in the explanatory variables because they have been regarded as likely sources of infection, or have shown a high sample correlation with the effective reproduction number. Because the definition of the effective reproduction number is the number of new cases per an infected person in the current population, real household expenditures in the explanatory variables are also on a per-household basis. In this paper, I use nationwide data because there is no publicly available household expenditure data for each prefecture at daily frequency. Because of data availability and the spread of mutant strains in 2021, the latter of which is likely to cause a structural change in the regression, the sample period for the regression is set to the period between February 15, 2020, and February 1, 2021.

In the regression model, the degree of daily infectious activities is assumed to be a linear function of the explanatory variables. The model incorporates time-varying coefficients through cross terms between the explanatory variables and time dummies. Then, infectious activities on each date contribute to the effective reproduction number from the next day to 14 days later, according to the distribution of incubation periods for new coronavirus. To minimize the number of parameters to estimate, I interpret the sample distribution of incubation periods in Japan as the probability distribution of incubation periods, and multiply the degree of infectious activities on each date with the probability of each incubation period as a weight. Thus, the effective reproduction number on each date is modeled as a weighted sum of lagged infectious activities over the past 14 days, excluding the current date. In this way, the regression model incorporates lagged explanatory variables without a need to create a new coefficient for each lag.

The regression model incorporates measurement error of the effective reproduction number, and also a latent AR(1) process for unobserved infectious activities on each date. To estimate this model, I use the Bayesian method with an uninformative, or improper, prior distribution for each parameter.

Estimated regression coefficients imply that, among real household expenditures in the explanatory variables, those for cafe and bar had larger effects on the effective reproduction number per value of spending than the other household expenditures during the sample period. Thus, a loss of aggregate demand is minimized if the effective reproduction number is lowered by restricting only household consumption of cafe and bar. This result is consistent with the fact that, up to the second state of emergency since the beginning of the pandemic, the government had been focusing on limiting the opening hours of bars and restaurants up to 8 p.m. in populated area, in order to curb infection through bar consumption at late night. In the third state of emergency from April 25, 2021, the government aims to cut the consumption of alcohol at bars and restaurants entirely in metropolitan area.

Given the regression model being in a linear reduced form, there remains a general concern on an omitted variable problem and specification error due to a possible non-linear relationship between the dependent variable and explanatory variables. To see if a bias in the regression model is small, I generate an out-of-sample prediction of the effective reproduction number for the case without any restriction on household consumption or mobility by inserting the 2019 data of explanatory variables into the estimated regression model. I will show that the simulated effective reproduction number falls within a range of epidemiologists' estimates of the basic reproduction number in China during an early phase of the pandemic between December 2019 and January 2020, when people in the country were yet to be fully adjusted to the pandemic.

Given this relatively good fit of an out-of-sample prediction from the estimated regression model, I use the model to run counterfactual simulations to obtain ballpark estimates of the effect of restricting household consumption of cafe and bar on the effective reproduction number. I find that even if a self-restraint on packaged domestic travels and an endogenous decline in mobility are taken into account, it will be necessary to cut household consumption of cafe and bar by more than 80% of the 2019 level, in order to keep the effective reproduction number below one on average. This result implies that the government must impose a severe volume restriction on not only bar consumption, but also cafe consumption, to contain the spread of new-coronavirus infection. If it is politically difficult to impose such a severe volume restriction on household consumption, then the government must shift the policy focus from a volume restriction to an intervention to reduce the infectiousness of household consumption of each type.

This paper is related to the literature on the relationship between mobility and new-coronavirus infection, such as Glaeser, Gorback, and Redding (2020) on U.S. data, and Watanabe and Yabu (2020), Kajitani and Hatayama (2021), and Kurita, Sugawara, and Ohkusa (2021) on Japanese data. Given a high correlation between mobility and household expenditures, the regression analysis in this paper can be interpreted as translating the infectious effect of mobility, which has been confirmed in the literature, into the infectious effect of real household expenditures. The latter measure is useful to discuss economic costs of policy interventions, because it is equivalent to the marginal economic cost to contain the spread of new-coronavirus infection in terms of a loss of aggregate demand.

This paper is also related to the large literature on the macroeconomic analysis of the new-coronavirus pandemic. Examples in Japan include Hamano, Katayama, and Kubota (2020), who endogenize a self-restraint on household consumption in an SIR-macro model, and

Fujii and Nakata (2021), who combine a reduce-form estimate of the effect of anti-infection social interventions on GDP with an SIR model.[1] While their top-down approaches are useful to endogenize GDP with the spread of infection, self-restricting behavior, and social interventions, this paper takes a bottom-up approach, providing reduced-form estimates of the effects of detailed household expenditures and mobility on new-coronavirus infection. This paper's approach is convenient when we discuss the marginal economic costs of detailed social interventions to contain the spread of new-coronavirus infection.

The remainder of this paper is organized as follows: Data sources and the selection of explanatory variables are described in section 2. The regression model is presented in section 3. The estimation of the regression model is reported in section 4. Simulations using the estimated regression model are described in section 5. Conclusions and discussion are in section 6.

2 Data

2.1 Data sources

Table 1 summarizes the sources of data used in this paper. The effective reproduction number published by Toyokeizai-Shinpo-Sha, a publisher in Japan, is the week-over-week gross rate of change in the number of new cases of new-coronavirus infection, raised to the power of 5/7, where 5 is the average generation time (i.e., the average number of days that it takes for an infected person to cause the next infection) and 7 is the number of days in a week. This simplified formula to estimate the effective reproduction number on the basis of reporting dates has been widely used in Japan to update the effective reproduction number real time daily.[2] In the Family Income and Expenditure Survey, daily data on nominal household

[1]For more examples of research in Japan, see the list collected by the Japanese Economic Association at `https://covid19.jeaweb.org/scientific.html`.

[2]For further discussion on the basis of this formula by Professor Hiroshi Nishiura of Kyoto University, a theoretical epidemiologist, in Japanese, see `https://github.com/contactmodel/COVID19-Japan-Reff` (accessed April 13, 2021).

expenditures are publicly available only for households with two or more members.

2.2 Sample correlation between the effective reproduction number and nominal household expenditures per household

Figure 1 plots the effective reproduction number and six types of nominal household expenditures per household: eating out for meals; cafe (including snack accompanying coffee and tea); bar (including meals accompanying alcoholic drink); lodging; domestic travel packages (i.e., bundles of lodging and transportation within the country); and clothing and footwear.[3] Household expenditures in the figure are 7-day backward moving averages, given the aforementioned formula for the effective reproduction number being an exponential function of the week-over-week gross rate of change in the number of new cases. The sample period starts from March 1, 2020, as the effective reproduction number from the data source is published only from that date.

The first five types of household expenditures have been regarded as potentially significant sources of infection. For example, the government shortened the opening hours of bars and restaurants in populated area during two states of emergency from April 7, 2020, to May 25, 2020, and from January 7, 2021, to March 21, 2021. Also, the government subsidized domestic traveling for sightseeing from July 22, 2020 to December 27, 2020, in order to make up for a loss of revenue for the tourism industry. This subsidy program was called a "Go-To-Travel" campaign.[4] There is a controversy over whether this campaign helped spreading new-coronavirus infection across the country.

In addition, Figure 1 includes clothing and footwear, because this item shows a high sample correlation with the effective reproduction number, as shown in Figure 2. In fact, clothing and footwear has a higher maximum cross correlation coefficient with the effective reproduction number than any other large category of nominal household expenditures, and

[3]Nominal household expenditure for foreign travel packages was negligible during the sample period.

[4]This campaign is supposed to resume in the future when the spread of new-coronavirus infection is contained, as of April 2021.

as high a maximum cross correlation coefficient as nominal household expenditure for bar (see Table 2).

Even though the purchase of clothing and footwear is not regarded as a potentially significant source of infection, the observed sample correlation implies that it is necessary to consider clothing and footwear as part of explanatory variables in the regression of the effective reproduction number on potential determinants, which will be shown in the next section. Not doing so may lead to a bias in the regression analysis, such that the infectious effect of shopping is estimated as part of the effect of eating out, because people often go to bars and restaurants after shopping on the high street.

2.3 Sample correlation between the effective reproduction number and mobility

Figure 3 plots the effective reproduction number and the six categories of the COVID-19 Community Mobility Reports from Google: retail_and_recreation; transit_stations; grocery_and_pharmacy; workplaces; parks; and residential. Among these, retail_and_recreation, transit_stations, grocery_and_pharmacy, and workplaces can cause human contacts outside families. retail_and_recreation, however, is closely correlated with nominal household expenditure per household on eating out for meals, as shown in Figure 4. To avoid a multicollinearity problem, I do not include retail_and_recreation as part of explanatory variables in the regression analysis shown in the next section. Among the remaining three categories of mobility data, transit_stations will be used as a general measure of mobility. This choice is due to convenience, as it allows me to use publicly available transportation data in 2019 for a substitute to Google data when I must use data in the period before mobility data from Google are available.

3 Regression model

3.1 Regression model and the definition of variables

Given the discussion described in the previous section, the log of the effective reproduction number is regressed on real household expenditures for eating out for meals, cafe, bar, lodging, domestic travel packages, and clothing and footwear, as well as transit_stations in the COVID-19 Community Mobility Reports from Google. Because the definition of the effective reproduction number is the number of new cases per an infected person in the current population, real household expenditures in the explanatory variables are also on a per-household basis.

Even though a low inflation rate in Japan makes the distinction between nominal and real household expenditures insignificant for most items, household expenditure for domestic travel packages is an exception, because the proportional subsidies during the "Go-To-Travel" campaign lowered nominal prices of traveling for sightseeing significantly during the campaign period. For this reason, I use real household expenditures for explanatory variables in the regression. Real household expenditures per household are computed by dividing nominal household expenditures per household by the corresponding categories of CPI for each, so that their unit is 100 yen in their 2020 average prices.[5]

[5]Because only monthly CPI is available, the value of CPI for each month is used for all dates within the same month. The CPI for eating out in general is used to convert nominal household expenditures for eating out for meals, cafe, and bar into real terms, because there is no separate CPI exactly corresponding for each. Because there is no corresponding CPI for domestic travel packages and because the CPI for lodging reflects not only the prices of independent lodging, but also the prices of lodging bundled with transportation in domestic travel packages, I use the CPI for lodging as a proxy to convert nominal household expenditure for domestic travel packages. On the other hand, perhaps because the Go-To-Travel campaign subsidized both costs of lodging and transportation costs, nominal household expenditure for domestic travel packages increased substantially during the campaign period, while that for lodging did not. To remove the effect of the Go-To-Travel campaign from the CPI for lodging, I interpolate the monthly CPI for lodging between July 2020 and January 2021 to convert nominal household expenditure for lodging in real terms. There is a corresponding CPI for clothing and footwear.

The form of the regression model is as follows:

$$\ln R_t = \sum_{s=0}^{6}(Z_{t-s} + \eta_{t-s}) \tag{1}$$

$$Z_t = \sum_{k=1}^{14} p_k V_{t-k} \tag{2}$$

$$\begin{aligned} V_t = {} & \alpha_0 + \alpha_1 D_{NY,t} + \alpha_2 D_{AH,t} + \sum_{j=0}^{2} \beta_j D_{SE,j,t} \\ & + \sum_{i=1}^{7} \left[\left(\gamma_i + \delta_i D_{AH,t} + \sum_{j=0}^{2} \phi_{j,i} D_{SE,j,t} \right) X_{i,t} \right] + e_t \end{aligned} \tag{3}$$

$$e_t = \rho e_{t-1} + \epsilon_t \tag{4}$$

where

$$\eta_t \sim N(0, \sigma_\eta^2), \tag{5}$$

$$\epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{6}$$

$$\gamma_i + \delta_i > 0, \quad \gamma_i + \delta_i + \phi_{j,i} > 0 \tag{7}$$

$$\delta_i < 0 \tag{8}$$

$$\rho \in (-1, 1) \tag{9}$$

The initial value of e_t in the estimation, denoted by e_0, is drawn from the unconditional probability distribution for e_t, given (4):

$$e_0 \sim N\left(0, \frac{\sigma_\epsilon^2}{1-\rho^2}\right) \tag{10}$$

The definition of variables is summarized in Table 3.

On the right-hand side of (1) is the sum of Z_{t-i} and η_{t-i} in the past 7 days, including the current date (i.e., for $i = 0, 1, ..., 6$), because the log of the effective reproduction number on the left-hand side is equivalent to the sum of the rate of change in the number of new cases in the past 7 days, multiplied by 5/7.

On the right-hand side of (2), p_k for $k = 1, 2, ..., 14$ is the sample distribution of incubation periods in Japan reported by Sugishita, Kurita, Sugawara, and Ohkusa (2020). See Figure 5 for the distribution. To compute the cumulative effect of lagged infectious events on new cases, Z_t, p_k is interpreted as the probability of the incubation period being k days. Then, p_k is multiplied to the degree of daily infectious events k days ago, i.e., V_{t-k}, for $k = 1, 2, ..., 14$, to measure the contribution from infectious events in k days ago for the rate of change in the number of new cases on each date. This use of the sample distribution of incubation periods makes it possible to incorporate a relatively long lag length (i.e., 14) without creating a new parameter for each lag. This is beneficial as the available sample period is less than a year (or 365 days).

In (3), the degree of infectious events on each date, V_t, is modeled as a linear function of real household expenditures per household and mobility in public transportation, which are denoted by $X_{i,t}$ for $i = 1, 2, ..., 7$. There are also time dummies for the year-end and new-year holiday period, $D_{NY,t}$, and for the period before the first state of emergency and the two states of emergency, $D_{SE,j,t}$ for $j = 0, 1, 2$, as well as a dummy for absolute humidity, $D_{AH,t}$. Through the cross terms between these dummies and $X_{i,t}$ for $i = 1, 2, ..., 7$, (3) implies that the infectious effects of economic activities are state-dependent. For the estimation of these effects, (7) imposes restrictions based on a prior expectation that in any state, economic activities measured by $X_{i,t}$ for $i = 1, 2, ..., 7$ spread new-coronavirus infection to some extent.

To compute $D_{AH,t}$ for each date, the dummy for absolute humidity no less than $9g/m^3$ for the capital of each prefecture is weighted by the population of the prefecture in 2019 and summed across prefectures to compute the population-weighted nationwide average of the dummies. The threshold level of absolute humidity at $9g/m^3$ for $D_{AH,t}$ is based on the fact that Nottmeyer and Sera (2021) report that the risk ratio of new cases of new-coronavirus infection over absolute humidity was non-linear, and peaked around $6 - 8g/m^3$ in their samples in England. $D_{AH,t}$ approximates such an effect of absolute humidity by a step function. See Figure 6 for the values of $D_{AH,t}$.

A caveat is that the risk ratio is just a sample correlation. Even though, to my knowledge, it is not clear whether there is established evidence for the biological effect of absolute humidity on the infectiousness of new coronavirus, (8) still imposes a negativity restriction on δ_i, i.e., the coefficient to the cross term between $D_{AH,t}$ and $X_{i,t}$, for $i = 1, 2, ..., 7$. This coefficient restriction is based on a prior expectation that at least the infectiousness of new coronavirus does not increase with absolute humidity.

3.2 Sample period

The sample period for the dependent variable is from March 6, 2020, to February 1, 2021. The beginning of the sample period is due to the availability of mobility data from Google.[6] The end of the sample period is set to include explanatory variables only up to January 2021 in the regression. This cap on the sample period is due to a concern on a possible spread of mutant strains in 2021. More specifically, the first report on the finding of a mutant strain from an airline passenger from abroad in Japan was on December 18, 2020.[7] By February 10, 2021, 108 cases of mutant strains had been found nationwide.[8] Also, the Tokyo Metropolitan Government started screening a sample of PCR-test results to detect mutant strains from December 2020, and found two cases of mutant strains from 1719 samples by January 29, 2021.[9] Thus, the spread of mutant strains was likely to be limited before the end of January 2021.

[6] The COVID-19 Community Mobility Reports from Google are available from February 15, 2020. There are 21 days between the first date of the dependent variable and that of the explanatory variables in the regression, because there are 14-day lags on the right-hand side of (2), and summation over 7 days on the right-hand side of (1).

[7] See `https://www.mhlw.go.jp/content/10900000/000764153.pdf` (accessed on April 14, 2021.)

[8] See `https://www3.nhk.or.jp/news/special/coronavirus/newvariant` (accessed on April 19, 2021.)

[9] See `https://www.metro.tokyo.lg.jp/tosei/hodohappyo/press/2021/01/30/01.html` (accessed on April 19, 2021.)

3.3 Clarification of possible biases

Before moving on, let me clarify possible biases in the regression model. Household expenditures and mobility may be contemporaneously affected by people's recognition of the current information set, including the current effective reproduction number. This channel, however, does not cause a simultaneity bias in the regression model because all explanatory variables lag the dependent variable due to incubation periods.

To discuss possible channels of an endogeneity bias further, Figure 7 shows a causal diagram. To understand the diagram, note that the effective reproduction number is determined by the product of three factors:[10]

- the rate of effective contact between an infected person and an unimmunized person;
- the probability of infection from an infected person to an unimmunized person per contact; and
- the average period of infection from an infected person.

The current effective reproduction number affects the rate of effective contact between an infected person and an unimmunized person in the future, because it determines the rate of increase in the immunized share of population. Thus, the current effective reproduction number can be a confounder for both the dependent variable and explanatory variables through this channel. However, given the immunized share of population remaining almost unchanged due to a relatively small number of total cases in Japan, the direct linkage between the current and future effective reproduction numbers through this channel is likely to be negligible during the sample period.[11]

[10] This decomposition is based on a non-technical summary of an SIR model by Suzuki and Nishiura (2020). Note that both the rate of effective contract and the probability of infection from an infected person are affected by social interventions.

[11] In contrast, Glaeser, Gorback, and Redding (2020) regress the log of the number of total cases of new-coronavirus infection per population or the log of new cases, rather than the effective reproduction number, on lagged mobility data across zip-code districts in New York City. Thus, even though their explanatory

Another concern is the possibility that people's expectations of future effective reproduction numbers affect people's decision making on their household expenditures and mobility today. Unlike economic variables, however, infection is never caused by a thought in human mind, including any type of expectation of future reproduction numbers, directly. The three determinants of the effective reproduction number are all physical factors, as listed above. Therefore, people's expectations do not cause an endogeneity bias if explanatory variables in the regression model include all types of physical behavior that affects the rate of human contacts among people.[12]

Nonetheless, it is difficult to include all possible physical factors for the effective reproduction number in the explanatory variables, given a small sample size and limited data availability. Thus, there remains a general concern on an omitted variable problem. Another general concern is specification error due to a possible non-linear relationship between the dependent variable and explanatory variables. To see if a bias in the model is small, I will report an out-of-sample prediction of the estimated regression model in a later section, instead of seeking a solution to all possible biases.[13] There, I will simulate the effective reproduction number for the case without any restriction on household consumption or mobility, and show that the simulation result is consistent with corresponding independent estimates reported by epidemiologists in literature.

variables lag the dependent variable, there is a significant concern on an endogeneity bias in their regression, because people may change their mobility after seeing the current number of total or new cases, and because the future numbers of total and new cases are determined by the product of the current number of infected people and the effective reproduction number, as implied by an SIR model. They resolve this issue by using indicators of the essential and the telecommuting share of workers in each district for instrumental variables. Likewise, Barro (2020) analyzes the effect of non-pharmaceutical interventions into the number of deaths across U.S. cities during the 1918-1920 Great Influenza Epidemic. Given non-pharmaceutical interventions by each municipal government were affected by the observed number of deaths, he uses the distance between each city and Boston, where the first case of infection was detected, for the instrumental variable.

[12] In other words, the sufficient state variables for the effective reproduction number are all physical variables. This feature of the effective reproduction number contrasts with usual economic analysis.

[13] Instrumental-variable estimation may be useful to correct an endogeneity bias due to an omitted variable problem, if an appropriate instrumental variable is known. Yet, it will not correct specification error in a linear regression model.

4 Estimation results

I apply the Bayesian method to estimate parameters in the regression. I set an uninformative, or improper, prior distribution for each parameter, that is, the density of the prior distribution of each set of parameter values is a constant, given the coefficient restrictions specified by (7) and (8). I use R ver. 4.0.3 (R Core Team 2020) and Rstan ver. 2.21.2 (Stan Development Team 2020) for estimation.[14]

Table 4 shows the posterior mean and the 95% credible interval of each parameter value. It is remarkable that the posterior means of γ_2 and γ_3, i.e., the coefficients to real household expenditures per household for cafe and bar, respectively, are much larger than those of γ_i for $i = 1, 4, 5, 6, 7$. This result implies that household expenditures for cafe and bar had larger effects on the effective reproduction number per value of spending than the other types of household expenditures in explanatory variables during the sample period. Thus, a loss of aggregate demand is minimized if the government aims to lower the effective reproduction number by restricting only cafe and bar consumption by households. This result is roughly consistent with the fact that, up to the second state of emergency since the beginning of the pandemic, the government had been focusing on limiting the opening hours of bars and restaurants up to 8 p.m. in populated area, in order to curb infection through bar consumption at late night. Also, the government aims to cut the consumption of alcohol at bars and restaurants entirely in metropolitan area in the third state of emergency from April 25, 2021.

The fitted value of the log of the effective reproduction number and also the residuals of the regression are shown in Figure 8. The fitted value deviates from the observed effective reproduction number substantially in the summer of 2020 and in November 2020. The bottom panels of the figure imply that these anomalies are mostly due to shocks to unobserved infectious events, rather than measurement error.

[14] The codes and data set for the estimation are available at `https://github.com/hajimetomura/R_HHexp`.

Even though the posterior mean of ϵ_t looks like having a serial correlation, the distributions of auto-correlation functions of residuals, i.e., η_t and ϵ_t, in the mcmc samples plotted in Figure 9 imply that serial correlation is mostly removed from residuals by the inclusion of an AR(1) process for unobserved infectious events, (4), in the regression.[15]

5 Simulations

5.1 Standard for policy evaluation

Hereafter, I simulate the quantitative effect of restricting cafe and bar consumption on the effective reproduction number, using the estimated coefficients of the regression model. For the measure to evaluate policy effects, I use the geometric mean of simulated effective reproduction numbers over a year. I highlight this indicator because if the effective reproduction number remains above one on average, then the number of new cases will exceed the finite capacity of medical services at some point in the future. Thus, given the prospect of population-wide availability of vaccinations in Japan looking still remote, the need to keep the effective reproduction number below one on average until then seems a socially agreeable target for the country. Even though choosing a year for the duration of the simulation period implies a pessimistic expectation that vaccinations will be widely available in the country only after a year, such a possibility is not entirely unrealistic in Japan as of April 2021. Using the annual mean of simulated effective reproduction numbers also allows to take into account the seasonality in household expenditures in simulations.

To clarify, the government may face a trade-off between new cases of new-coronavirus infection (or deaths) and a measure of economic activities such as GDP, if it aims to stabilize the effective reproduction number at some specific level between 0 and 1 until the arrival of vaccinations for a sufficiently large part of the population, because the total number of

[15] In mcmc sampling, the value of ϵ_t is simulated to compute the likelihood of the value of η_t, i.e., the residual of the observation equation, (1). As a result, the auto-correlation function of ϵ_t is smooth around 0, whereas that for η_t is more fluctuating, as shown in Figure 9.

deaths due to new coronavirus will be lower as the targeted value of the effective reproduction number is set closer to zero. This question is beyond the scope of this paper.

5.2 Benchmark simulation with hypothetical 2019 data

To set a benchmark, I first simulate the effective reproduction number with the 2019 data of explanatory variables, which can be interpreted as a hypothetical case of no restriction on household consumption or mobility. Because the COVID-19 Community Mobility Reports from Google does not exist for 2019, I create an index of mobility in public transportation for 2019 by dividing the monthly average of railway passengers in each month by the monthly average in January 2020. This indexation is consistent with the feature of the COVID-19 Community Mobility Reports such that each type of mobility data in the reports are expressed in the form of the rate of change from the average over the period between January 3 and February 6 in 2020. Because only the monthly averages of railway passengers are publicly available, I simply use the monthly average in each month for the daily value on each date within the same month. This substitution can be justified by a high correlation between transit_stations in the COVID-19 Community Mobility Reports from Google and the monthly average of railway passengers in 2020, as shown in Figure 10.

Using the 2019 data, I simulate the effective reproduction number for 365 days from March 6, which coincides with the first date of the effective reproduction number in the regression model with 2020-2021 data.[16] To simulate the effective reproduction number for a year, I connect the year end of the 2019 data with the new year data on January 1, 2019, so that the 2019 data loop as hypothetical data without any restriction on household consumption or mobility. Figures 11 and 12 compare the 2019 data of explanatory variables with the 2020-21 data used for estimation of the regression model.

To make comparison between the simulated and the observed values of the effective

[16] For the simulation, I use the data of explanatory variables from February 14, 2019, i.e., one day before the sample period of explanatory variables for the estimation, because 2020 is a leap year.

reproduction number, I only change the values of real household expenditures and mobility in public transportation to the 2019 data in the simulation. I keep using the 2020-21 data for absolute humidity (i.e., $D_{AH,t}$) as well as the dummy for the year-end and new-year holiday period (i.e., $D_{NY,t}$). I set zero to all dummies related to the states of emergency (i.e., $D_{SE,j,t} = 0$ for $j = 0, 1, 2$).

Figure 13 plots the posterior mean and the 95% credible interval of $\ln R_t$ in the simulation with hypothetical 2019 data, along with the observed and the fitted value of $\ln R_t$ for 2020-2021 from March 6, 2020, over a year. The figure indicates that without any restriction on household consumption or mobility, the effective reproduction number would rise around the end of the fiscal year (i.e., the end of March); after the Golden Week holiday period in early May; and in November and December.[17]

Table 5 summarizes the posterior distribution of annual means of $\ln R_t$ in the simulation with hypothetical 2019 data. Because $\ln R_t$ is the log of the effective reproduction number, the annual mean of $\ln R_t$ corresponds to the log of the geometric annual mean of the effective reproduction number. To make it easy to interpret the simulation result, the table shows the exponential value of each figure in the parenthesis below the figure.

Because the 2019 data in the simulation are used as hypothetical data of real household expenditures and mobility without any policy intervention or self-restraint, the geometric annual mean of the effective reproduction number generated by the simulation with the hypothetical 2019 data is comparable with a hypothetical basic reproduction number (i.e., the average number of new cases per an infected person in a population where everyone is susceptible to infection). Indeed, the simulation result shown in Table 5 is largely consistent with the range of existing estimates of the basic reproduction number during an early phase of the pandemic in China between December 2019 and January 2020, when people in the country were yet to be fully adjusted to the pandemic. The range was between 1.4 and

[17]A caveat is that the Golden Week holiday period in 2019 lasted for 10 days, which was longer than usual. Thus, an increase in household consumption during the Golden Week in 2019 could be higher than that in a regular year. I thank Hiroshi Fujiki for pointing out this anomaly in 2019.

3.5 (see Imai, et al., 2020). This relatively good fit of an out-of-sample prediction of the regression model provides a support for using the model for counterfactual simulations. A similar result can be obtained even if the level of the effective reproduction number, R_t, is used for the dependent variable in the regression model, instead of $\ln R_t$.[18]

5.3 Endogenizing the response of mobility in public transportation to an exogenous restriction on household consumption

As described above, Table 4 implies that a loss of aggregate demand is minimized if the effective reproduction number is lowered by restricting only cafe and bar consumption by households. Hereafter, I simulate the quantitative effects of percentage reductions of cafe and bar consumption by households compared to the 2019 level. For this simulation, I will consider four scenarios that are characterized by two factors: whether the self-restraint on packaged domestic travels observed in the first half of 2020 is assumed to continue or not; and whether an endogenous response of mobility in public transportation is taken into account or not. Table 6 summarizes the characteristics of the four scenarios.

To endogenize the response of mobility in public transportation to a restriction on household consumption, I regress transit_stations in the COVID-19 Community Mobility Reports from Google (i.e., $X_{7,t}$) on real household expenditures per household (i.e., $X_{i,t}$ for $i = 1, 2, ..., 6$) among the explanatory variables of the regression model for the effective reproduction number. I also include the following time dummies as part of explanatory variables to capture the seasonality of mobility: holidays, including weekends and the year-end and new-year holiday period from December 29-January 3; each of the two states of emergency from April 7, 2020, to May 25, 2020 and from January 7, 2021, to March 21, 2021; and December, which is due to a change in the relationship between bar consumption and mobility due to year-end parties. I estimate the regression coefficients by OLS. The sample period is from February 15, 2020, to January 31, 2021, which is the same as the sample

[18] See appendix for the results of estimation and simulations in this case.

period of explanatory variables in the estimation of the regression model for the effective reproduction number.

Among the explanatory variables, eating out for meals (i.e., $X_{1,t}$) has a statistically insignificant coefficient. Table 7 reports the OLS estimate of the regression of transit_stations without eating out for meals in the explanatory variables.

To see the fit of this regression by an out-of-sample prediction, Figure 14 plots the ratio of the monthly average of railway passengers in 2019 to the January 2020 average, and the monthly average of daily fitted values generated by applying estimated regression coefficients shown in Table 7 to the 2019 data of the explanatory variables.[19] The difference between the means of the two series in the figure can be interpreted as a time fixed effect. The figure shows that the fitted values largely replicate the observed pattern of time variations in the number of railway passengers in 2019, such that the number of railway passengers drops significantly in February and December, while fluctuating around a stable level from March to November. Hereafter, I use the regression coefficients shown in Table 7 when I endogenize the response of mobility in public transportation to an exogenous restriction on household consumption.

5.4 Quantitative effect of restricting household consumption on the effective reproduction number

Now, let me describe the results of simulations from 1 to 4. In simulation 1, only cafe and bar consumption is restricted in order to minimize a loss of aggregate demand to reduce the effective reproduction number, while an endogenous response of mobility in transportation is taken into account. Table 8 shows the posterior distribution of annual means of $\ln R_t$ for a range of percentage reductions of cafe and bar consumption by households from the 2019 level. To interpret figures in the table, note that zero corresponds to the case in which the geometric annual mean of the effective reproduction number is 1. The posterior means

[19]Time dummies for the two states of emergency are set to zero in the 2019 data.

in the table imply that it will be necessary to cut 95-100% of cafe and bar consumption by households compared to the 2019 level, if the government aims to stabilize the effective reproduction number below 1 on average throughout a year. Figure 15 shows the effective reproduction number, hypothetical real household expenditures per household for cafe and bar, and endogenized mobility in simulation 1, when the restriction on cafe and bar consumption by households is sufficiently large to make the posterior mean of the annual mean of $\ln R_t$ below zero.

Even though simulation 1 takes into account the endogenous response of mobility to a restriction on cafe and bar consumption by households, the regression of mobility in public transportation summarized in Table 7 is likely to suffer some degree of endogeneity bias. For robustness check, simulation 2 assumes only a restriction on cafe and bar consumption while keeping mobility in public transportation at its 2019 level as shown in Figure 12. Table 9 shows the posterior distribution of annual means of $\ln R_t$ in simulation 2. The table implies that without an endogenous response of mobility, the government cannot keep the effective reproduction number below 1 on average by restricting only cafe and bar consumption by households.

In simulations 1 and 2, it is assumed that the government aims to restrict only cafe and bar consumption by households in order to minimize a loss of aggregate demand for the reduction of the effective reproduction number. Therefore, real household expenditure for domestic travel packages is set to be as high as in 2019. This assumption may be unrealistic, because there has been a self-restraint on packaged domestic travels in 2020, except for the Go-To-Travel campaign period between July 22 and December 27 in 2020, as shown in Figure 11. To take into account this observation, it is assumed in simulations 3 and 4 that real household expenditure per household for domestic travel packages will be as low as the average in the period between the end of the first state of emergency and the beginning of the Go-To-Travel campaign period, i.e., from May 26, 2020, to July 21, 2020. Mobility in public transportation is endogeneized in simulation 3, while it is not in simulation 4, as described in Table 6.

Tables 10 and 11 show the posterior distributions of annual means of $\ln R_t$ in simulations 3 and 4, respectively. Table 10 implies that even with a self-restraint on packaged domestic travels, the government would need to cut 80-85% of cafe and bar consumption by households to stabilize the effective reproduction number below 1 on average throughout a year. If mobility does not decline endogenously, the necessary percentage reduction of cafe and bar consumption by households would rise to 95-100%. Figure 16 shows the effective reproduction number, hypothetical real household expenditures per household for cafe and bar, and endogenized mobility in simulation 3, when the restriction on bar and cafe consumption by households is sufficiently large to make the posterior mean of the annual means of $\ln R_t$ below one.[20]

6 Conclusions

In this paper, I regress the log of the estimate of the effective reproduction number based on reporting dates on a set of real household expenditures per household that are regarded as infectious, and also a measure of mobility in public transportation, using publicly available daily nationwide data in Japan. The estimation result indicates that a loss of aggregate demand will be minimized if the effective reproduction number is lowered by cutting only household consumption of cafe and bar. This result is largely consistent with the government policy up to the third state of emergency that began on April 25, 2021, which have been limiting the opening hours of bars and restraints up to 8 p.m. or cutting the consumption of alcohol at bars and restaurants entirely. The simulation results, however, indicate that even if a self-restraint on packaged domestic travels and an endogenous decline in mobility are

[20]Note that the hypothetical values of real household expenditures per household are the same between simulations 3 and 4. In simulation 4, mobility in public transportation takes the same value as 2019 data shown in Figure 12 for each date.

taken into account, it would be necessary to cut more than 85% of household consumption of both cafe and bar compared to the 2019 level, in order to keep the effective reproduction number below one on average.

One possible interpretation of this result is that a much more severe restriction on the opening hours of bars and restaurants than the aforementioned government policy will be necessary to prevent the explosion of new-coronavirus infection until vaccinations become widely available across the population. Another interpretation is that given the difficulty to impose such a severe restriction on bars and restaurants in reality, it is necessary to shift the policy focus from a volume restriction, such as shortening the opening hours of bars and restaurants, to an increase in the quality of efforts to reduce the infectiousness of each economic activity, which is measured by the coefficients of the regression model in this paper.

In addition, the comparison between simulations with and without endogenized mobility indicates that the effect of restricting household consumption of cafe and bar depends on how much mobility will decline in response to such a restriction. This result implies that even though explanatory variables in the regression model incorporate a representative set of infectious household expenditures, there still remains an infectious effect of mobility separately from those household expenditures. It is an important issue to figure out unidentified economic activities behind this effect of mobility.

In this paper, I do not simulate the effect of restricting the other types of household expenditures than cafe and bar consumption, such as eating out for meals, traveling, and apparel shopping. This is because there seems still room for improvement in the efforts to reduce infectiousness of these household expenditures, such as reducing oral conversation while eating, moving, and shopping outside.

Likewise, I do not simulate the effect of an exogenous restriction on mobility on the effective reproduction number. This is because such an intervention is likely to change household consumption endogenously. Thus, the economic cost of such a restriction cannot be measured by the reduced-form regression in this paper. Also, I do not take into account the substitution effect of a restriction on household consumption of cafe and bar that may increase other types of household consumption in the simulation exercises. With these reservations, this paper provides ballpark estimates of the effect of a cost-effective intervention in household consumption on the effective reproduction number.

References

[1] Barro, Robert J., 2020. "Non-Pharmaceutical Interventions and Mortality in U.S. Cities during the Great Influenza Pandemic, 1918-1919," NBER Working Paper 27049, National Bureau of Economic Research. https://doi.org/10.3386/w27049 (accessed April 28, 2021.)

[2] Fujii, Daisuke, and Taisuke Nakata, 2021. "Covid-19 and Output in Japan," RIETI Working Paper 21-E-004, Research Institute of Economy, Trade and Industry. https://www.rieti.go.jp/jp/publications/dp/21e004.pdf (accessed April 28, 2021.)

[3] Glaeser, Edward L., Caitlin Gorback, and Stephen J. Redding, 2020. "JUE Insight: How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities." Journal of Urban Economics. https://doi.org/10.1016/j.jue.2020.103292 (accessed April 28, 2021.)

[4] Hamano, Masashige, Munechika Katayama, and So Kubota (2020) "COVID-19 Misperception and Macroeconomy," WINPEC Working Paper Series No. E2016, Waseda University. https://www.waseda.jp/fpse/winpec/assets/uploads/2020/11/E2016_20201102-1version.pdf (accessed April 28, 2021.)

[5] Imai N, et al., 2020. Report 3 : Transmissibility of 2019-nCoV. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-01-25-COVID19-Report-3.pdf (accessed April 28, 2021.)

[6] Kajitani, Yoshio, and Michinori Hatayama, 2021. "Explaining the effective reproduction number of COVID-19 through mobility and enterprise statistics: Evidence from the first wave in Japan," PLoS ONE 16(3): e0247186. https://doi.org/10.1371/journal.pone.0247186 (accessed April 28, 2021.)

[7] Kurita, Junko, Tamie Sugawara, and Yasushi Ohkusa, 2021. "Effects of climate conditions, mobility trends, and countermeasures on the COVID-19 outbreak in Japan," Medrxiv, https://doi.org/10.1101/2020.12.29.20248977 (accessed April 28, 2021.)

[8] Nottmeyer, Luise N., and Francesco Sera, 2021. "Influence of temperature, and of relative and absolute humidity on COVID-19 incidence in England - A multi-city time-series study," Environmental Research, 196:1109-77. https://doi.org/10.1016/j.envres.2021.110977 (accessed April 28, 2021.)

[9] R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (accessed April 28, 2021.)

[10] Stan Development Team, 2020. Stan Modeling Language Users Guide and Reference Manual, VERSION 2.21.2. https://mc-stan.org (accessed April 28, 2021.)

[11] Sugishita, Yoshiyuki, Junko Kurita, Tamie Sugawara, and Yasushi Ohkusa, 2020. Effects of voluntary event cancellation and school closure as countermeasures against COVID-19 outbreak in Japan. PLoS ONE 15(12): e0239455. https://doi.org/10.1371/journal.pone.0239455 (accessed April 28, 2021.)

[12] Suzuki, Ayako, and Hiroshi Nishiura, 2020. "COVID-19. Topics: IV. Mathematical modeling and control of infectious diseases," The Journal of the Japanese Society of Internal Medicine, 109:2276-2280. https://www.naika.or.jp/jsim_wp/wp-content/uploads/2020/11/nichinaishi-109-11-article_4.pdf (accessed April 28, 2021.)

[13] Watanabe, Tsutomu, and Tomoyasu Yabu, 2020. "Japan's Voluntary Lockdown", Covid Economics 46(1), pp. 1-31. https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0 (accessed May 7, 2021.)

A Estimation of the regression model with R_t for the dependent variable

In this appendix, I report the estimation result of the regression in which R_t is used for the dependent variable instead of $\ln R_t$. Given the difficulty to obtain the convergence of mcmc, I tighten the coefficient restriction for ρ from $\rho \in (-1, -1)$ to $\rho \in (0, -1)$, given the 95% credible interval of ρ in the original regression model with $\ln R_t$ being strictly positive, as shown in Table 9. The estimation result is shown in Table 12. The simulation result corresponding to Table 5 is shown in Table 13.

Table 1: Data sources

Data	Level	Frequency	Source
Effective reproduction number	Nationwide	Daily	Toyokeizai-Shinpo-Sha
Nominal household expenditures per household	Nationwide	Daily	Households with two or more members, Family Income and Expenditure Survey, Ministry of Internal Affairs and Communications
Consumer Price Index (CPI)	Nationwide	Monthly	Ministry of Internal Affairs and Communications
Mobility in public transportation	Nationwide	Daily	transit_stations, COVID-19 Community Mobility Reports, Google
Temperature, Relative humidity	Prefectural	Daily	Japan Meteorological Agency
Populations	Prefectural	Annual	Population estimates, Ministry of Internal Affairs and Communications
Railway passengers	Nationwide	Monthly	Statistical Survey on Railway Transport, Ministry of Land, Infrastructure, Transport and Tourism
Sample distribution of incubation periods	Nationwide	–	Sugishita, Kurita, Sugawara, and Ohkusa (2020)

Table 2: Cross correlation coefficients between the effective reproduction number and 7-day moving averages of nominal household expenditures of large categories per household

	Maximum cross correlation coefficient	Corresponding lag of nominal household expenditures
Food	0.22	10
Housing	0.16	16
Fuel, light and water charges	-0.01	9
Furniture and household utensils	0.26	10
Clothing and footwear	0.62	12
Medical care	0.20	1
Transportation and communication	0.31	10
Education	0.44	5
Culture and recreation	0.38	9
Other consumption expenditures	0.48	8
Bar	0.66	9

Notes: The table shows the maximum cross correlation coefficients between the contemporaneous effective reproduction number and lagged 7-day backward moving averages of nominal household expenditures per household. The sample period is from March 1, 2020, to February 28, 2021, as the effective reproduction number is available only from March 1, 2020.

Table 3: Definition of variables

R_t	Effective reproduction number
$X_{1,t}$	Real household expenditure per household on eating out for meals
$X_{2,t}$	Real household expenditure per household for cafe
$X_{3,t}$	Real household expenditure per household for bar
$X_{4,t}$	Real household expenditure per household for lodging
$X_{5,t}$	Real household expenditure per household for domestic travel packages
$X_{6,t}$	Real household expenditure per household for clothing and footwear
$X_{7,t}$	transit_stations in the COVID-19 Community Mobility Reports for Japan, nationwide
$D_{SE,0,t}$	Time dummy for the period before the first state of emergency (- 2020/4/6)
$D_{SE,1,t}$	Time dummy for the first state of emergency (2020/4/7-2020/5/25)
$D_{SE,2,t}$	Time dummy for the second state of emergency (2021/1/7-2021/3/21)
$D_{NY,t}$	Time dummy for Dec. 29-Jan. 3.
$D_{AH,t}$	Population-weighted average of the dummy for absolute temperature no less than $9g/m^3$ across the capitals of prefectures.
p_k	A sample distribution of incubation periods in Japan.
V_t	Degree of daily infectious events.
Z_t	Cumulative effect of lagged infectious events on new cases of new-coronavirus infection.
e_t	Unobserved infectious events.
ϵ_t	Shocks to unobserved infectious events.
η_t	Measurement error.

Notes: The effective reproduction number is the week-over-week gross rate of change in the number of new cases, raised to the power of 5/7. The unit of each type of real household expenditure per household is 100 yen in the 2020 average price. To compute $D_{AH,t}$ for each date, the dummy for absolute temperature no less than $9g/m^3$ is constructed for the capital of each prefecture, weighted by the population estimate for the prefecture in 2019, and then summed across prefectures to compute the population-weighted average of the dummies.

Table 4: Estimated regression coefficients

	Posterior mean	2.5%	97.5%		Posterior mean	2.5%	97.5%
α_0	-0.083	-0.186	0.014	ϕ_{01}	-0.001	-0.023	0.023
α_1	0.054	0.002	0.154	ϕ_{02}	-0.003	-0.353	0.355
α_2	-0.018	-0.061	-0.001	ϕ_{03}	-0.024	-0.209	0.151
β_0	-0.100	-0.289	0.072	ϕ_{04}	0.010	-0.068	0.118
β_1	0.073	-0.212	0.384	ϕ_{05}	0.008	-0.047	0.081
β_2	0.220	-0.301	0.807	ϕ_{06}	0.032	-0.015	0.088
γ_1	0.012	0.001	0.032	ϕ_{07}	-0.000	-0.002	0.002
γ_2	0.187	0.025	0.515	ϕ_{11}	0.021	-0.016	0.086
γ_3	0.108	0.013	0.280	ϕ_{12}	0.425	-0.237	1.759
γ_4	0.047	0.007	0.112	ϕ_{13}	0.399	-0.099	1.173
γ_5	0.031	0.004	0.084	ϕ_{14}	0.614	-0.002	1.678
γ_6	0.018	0.002	0.041	ϕ_{15}	0.993	0.077	2.123
γ_7	0.002	0.000	0.005	ϕ_{16}	0.013	-0.025	0.076
δ_1	-0.002	-0.009	-0.000	ϕ_{17}	0.004	-0.001	0.010
δ_2	-0.051	-0.190	-0.001	ϕ_{21}	0.033	-0.014	0.130
δ_3	-0.025	-0.092	-0.001	ϕ_{22}	2.080	-0.033	5.749
δ_4	-0.016	-0.057	-0.000	ϕ_{23}	1.038	-0.050	2.976
δ_5	-0.013	-0.046	-0.000	ϕ_{24}	0.195	-0.036	0.670
δ_6	-0.004	-0.013	-0.000	ϕ_{25}	0.649	0.019	1.585
δ_7	-0.001	-0.003	-0.000	ϕ_{26}	0.029	-0.022	0.118
ρ	0.743	0.346	0.959	ϕ_{27}	0.020	0.002	0.043
σ_η	0.027	0.024	0.029				
σ_ϵ	0.050	0.030	0.088				

Notes: "2.5%" and "97.5%" indicate the percentiles of mcmc samples. The sample period for the dependent variable is from March 6, 2020, to February 1, 2021. The number of observations is 333. The prior distribution is an improper distribution for each parameter.

Table 5: Posterior distribution of annual means of $\ln R_t$ in the simulation with hypothetical 2019 data

	Posterior mean	2.5% percentile	97.5% percentile
Annual mean of $\ln R_t$	0.94	0.49	1.57
(Corresponding geometric annual mean of R_t)	(2.57)	(1.63)	(4.81)

Note: Each cell shows the posterior mean or a percentile of annual means of $\ln R_t$ simulated by inserting the hypothetical 2019 data of real household expenditures and mobility in public transportation into the regression model for the effective reproduction number. In the parenthesis below each figure is the exponential value of the figure, which corresponds to the geometric annual mean of the effective reproduction number implied by the figure.

Table 6: Four scenarios for the simulations of restrictions on cafe and bar consumption

	Self-restraint on packaged domestic travels	Endogenized mobility in public transportation
Simulation 1	No	Yes
Simulation 2	No	No
Simulation 3	Yes	Yes
Simulation 4	Yes	No

Table 7: Regression of mobility in public transportation on real household expenditures per household

	OLS estimate	Standard deviation	t value
Intercept	0.653	0.014	45.66
$X_{2,t}$	0.608	0.089	6.82
$X_{3,t}$	0.211	0.026	7.90
$X_{4,t}$	-0.045	0.014	-3.12
$X_{5,t}$	0.038	0.010	3.75
$X_{6,t}$	0.011	0.004	2.66
Dummy for holidays	-0.147	0.01	-12.97
Dummy for the first state of emergency	-0.141	0.013	-10.51
Dummy for the second state of emergency	-0.063	0.014	-4.54
Dummy for December	0.025	0.025	1.00
$X_{3,t}$*(Dummy for December)	-0.179	0.138	-1.29

Dependent variable: $1 + X_{7,t}/100$.

R^2: 0.77; adj. R^2: 0.76.

Sample period: Februrary 15, 2020, - January 31, 2021.

Table 8: Posterior distribution of annual means of $\ln R_t$ in simulation 1

Degree of % reduction of cafe and bar consumption compared to the 2019 level	Posterior mean	2.5% percentile	97.5% percentile
50%	0.40	0.11	0.76
55%	0.35	0.06	0.73
60%	0.31	0.00	0.69
65%	0.27	-0.04	0.65
70%	0.23	-0.10	0.62
75%	0.19	-0.16	0.58
80%	0.14	-0.23	0.54
85%	0.10	-0.30	0.51
90%	0.06	-0.37	0.48
95%	0.02	-0.44	0.45
100%	-0.02	-0.51	0.42

Notes: In simulation 1, no self-restraint on packaged domestic travels is considered, whereas mobility in public transportation is endogenized by the regression shown in Table 7. Each figure is the annual mean of $\ln R_t$ for an exogenous percentage reduction of cafe and bar consumption by households compared to the 2019 level in the first column. For each figure, 0 corresponds to the case in which the geometric annual mean of the effective reproduction number is 1.

Table 9: Posterior distribution of annual means of $\ln R_t$ in simulation 2

Degree of % reduction of cafe and bar consumption compared to the 2019 level	Posterior mean	2.5% percentile	97.5% percentile
50%	0.39	0.10	0.75
55%	0.36	0.06	0.73
60%	0.32	0.01	0.70
65%	0.29	-0.02	0.68
70%	0.26	-0.06	0.65
75%	0.23	-0.11	0.62
80%	0.20	-0.16	0.61
85%	0.17	-0.22	0.59
90%	0.14	-0.28	0.57
95%	0.11	-0.34	0.55
100%	0.08	-0.39	0.53

Notes: In simulation 2, no self-restraint on packaged domestic travels or endogenous response of mobility in public transportation is considered. Each figure is the annual mean of $\ln R_t$ for an exogenous percentage reduction of cafe and bar consumption by households compared to the 2019 level in the first column. For each figure, 0 corresponds to the case in which the geometric annual mean of the effective reproduction number is 1.

Table 10: Posterior distribution of annual means of $\ln R_t$ in simulation 3

Degree of % reduction of cafe and bar consumption compared to the 2019 level	Posterior mean	2.5% percentile	97.5% percentile
50%	0.29	-0.00	0.62
55%	0.24	-0.06	0.58
60%	0.20	-0.11	0.53
65%	0.16	-0.16	0.49
70%	0.12	-0.22	0.46
75%	0.08	-0.29	0.42
80%	0.03	-0.35	0.39
85%	-0.00	-0.41	0.36
90%	-0.04	-0.48	0.34
95%	-0.08	-0.55	0.31
100%	-0.13	-0.63	0.28

Notes: In simulation 3, both a self-restraint on packaged domestic travels and an endogenous response of mobility in public transportation are considered. Each figure is the annual mean of $\ln R_t$ for an exogenous percentage reduction of cafe and bar consumption by households compared to the 2019 level in the first column. For each figure, 0 corresponds to the case in which the geometric annual mean of the effective reproduction number is 1.

Table 11: Posterior distribution of annual means of $\ln R_t$ in simulation 4

Degree of % reduction of cafe and bar consumption compared to the 2019 level	Posterior mean	2.5% percentile	97.5% percentile
50%	0.30	0.00	0.63
55%	0.26	-0.04	0.60
60%	0.23	-0.09	0.56
65%	0.20	-0.13	0.53
70%	0.17	-0.17	0.51
75%	0.14	-0.22	0.49
80%	0.11	-0.26	0.47
85%	0.08	-0.31	0.45
90%	0.05	-0.37	0.44
95%	0.02	-0.42	0.42
100%	-0.00	-0.48	0.40

Notes: In simulation 4, a self-restraint on packaged domestic travels is considered while mobility in public transportation is not endogenized. Each figure is the annual mean of $\ln R_t$ for an exogenous percentage reduction of cafe and bar consumption by households compared to the 2019 level in the first column. For each figure, 0 corresponds to the case in which the geometric annual mean of the effective reproduction number is 1.

Table 12: Estimated regression coefficients of an alternative regression model

	Posterior mean	2.5%	97.5%		Posterior mean	2.5%	97.5%
α_0	0.058	-0.132	0.254	ϕ_{01}	-0.002	-0.033	0.031
α_1	0.094	0.004	0.261	ϕ_{02}	-0.076	-0.653	0.447
α_2	-0.032	-0.111	-0.001	ϕ_{03}	-0.058	-0.297	0.136
β_0	-0.323	-0.640	-0.000	ϕ_{04}	0.011	-0.095	0.150
β_1	-0.083	-0.528	0.308	ϕ_{05}	0.017	-0.065	0.129
β_2	0.640	-0.215	1.539	ϕ_{06}	0.105	0.025	0.190
γ_1	0.017	0.002	0.046	ϕ_{07}	-0.001	-0.005	0.003
γ_2	0.316	0.037	0.880	ϕ_{11}	0.034	-0.021	0.135
γ_3	0.145	0.017	0.374	ϕ_{12}	0.643	-0.387	2.585
γ_4	0.062	0.008	0.153	ϕ_{13}	0.403	-0.159	1.376
γ_5	0.051	0.006	0.133	ϕ_{14}	0.608	-0.025	1.965
γ_6	0.023	0.003	0.058	ϕ_{15}	1.074	0.040	2.722
γ_7	0.003	0.001	0.008	ϕ_{16}	0.023	-0.032	0.116
δ_1	-0.004	-0.014	-0.000	ϕ_{17}	0.001	-0.004	0.008
δ_2	-0.079	-0.295	-0.002	ϕ_{21}	0.034	-0.022	0.149
δ_3	-0.031	-0.112	-0.001	ϕ_{22}	1.968	-0.243	6.378
δ_4	-0.022	-0.079	-0.001	ϕ_{23}	1.359	-0.090	4.161
δ_5	-0.023	-0.086	-0.001	ϕ_{24}	0.408	-0.031	1.241
δ_6	-0.006	-0.023	-0.000	ϕ_{25}	0.558	-0.005	1.659
δ_7	-0.001	-0.004	-0.000	ϕ_{26}	0.036	-0.028	0.156
ρ	0.779	0.030	0.992	ϕ_{27}	0.033	0.005	0.066
σ_η	0.041	0.033	0.045				
σ_ϵ	0.070	0.038	0.192				

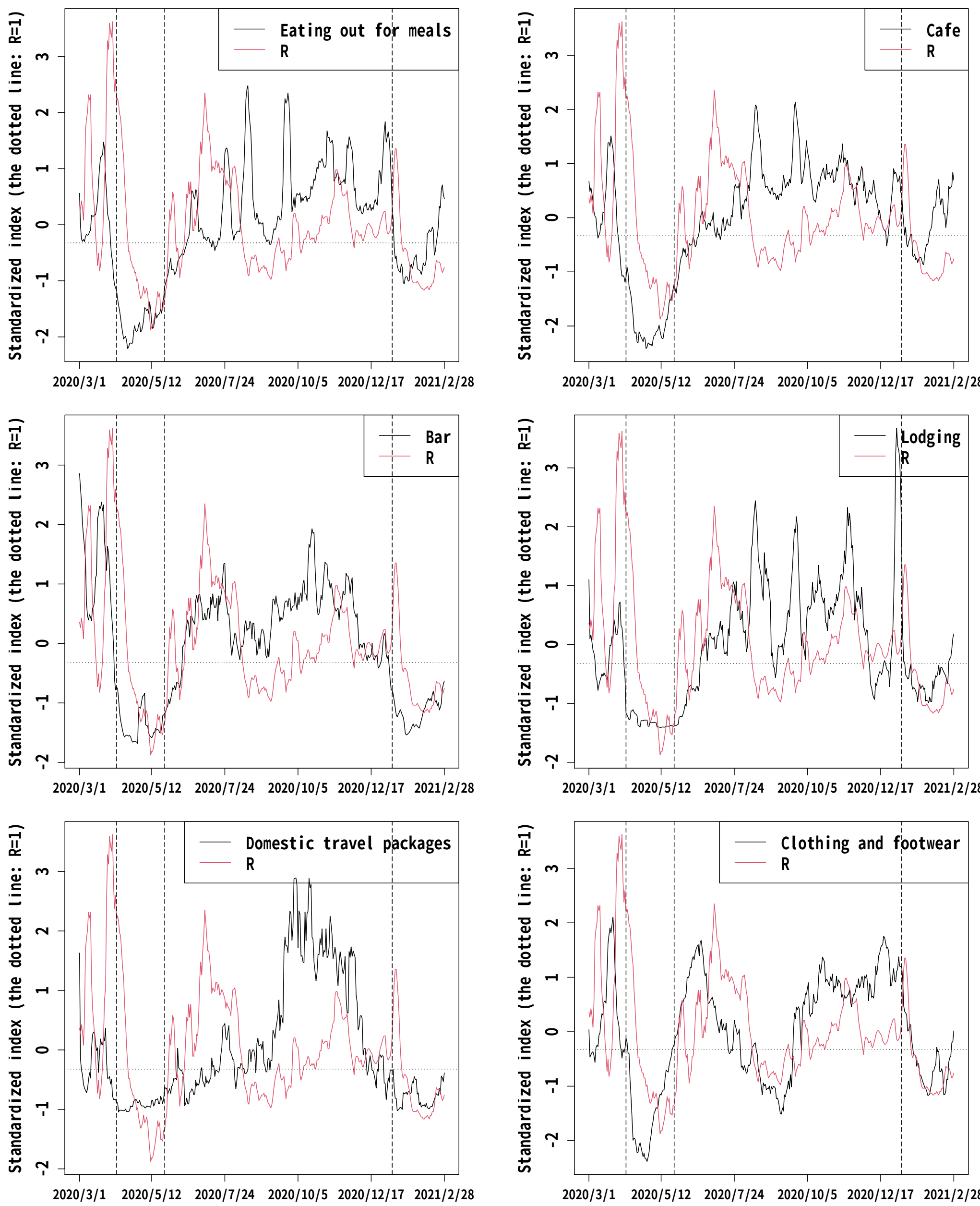
Notes: The alternative regression model has R_t for the dependent variable, instead of $\ln R_t$, in (1). "2.5%" and "97.5%" indicate the percentiles of mcmc samples. The sample period for the dependent variable is from March 6, 2020 to February 1, 2021. The number of observations is 333. The prior distribution is an improper distribution for each parameter.

Table 13: Posterior distribution of annual means of R_t in the simulation of an alternative regression model with hypothetical 2019 data

	Posterior mean	2.5% percentile	97.5% percentile
Annual mean of R_t	2.61	1.54	4.56

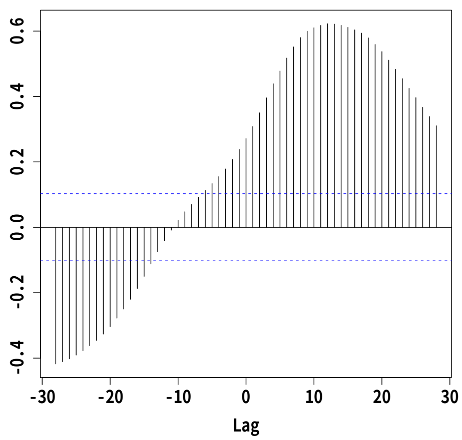
Note: The alternative regression model has R_t for the dependent variable, instead of $\ln R_t$, in (1). Each cell shows the posterior mean or a percentile of annual means of R_t simulated by inserting hypothetical 2019 data of real household expenditures and mobility in public transportation into the regression model.

Figure 1: Effective reproduction number and 7-day moving averages of nominal household expenditures per household



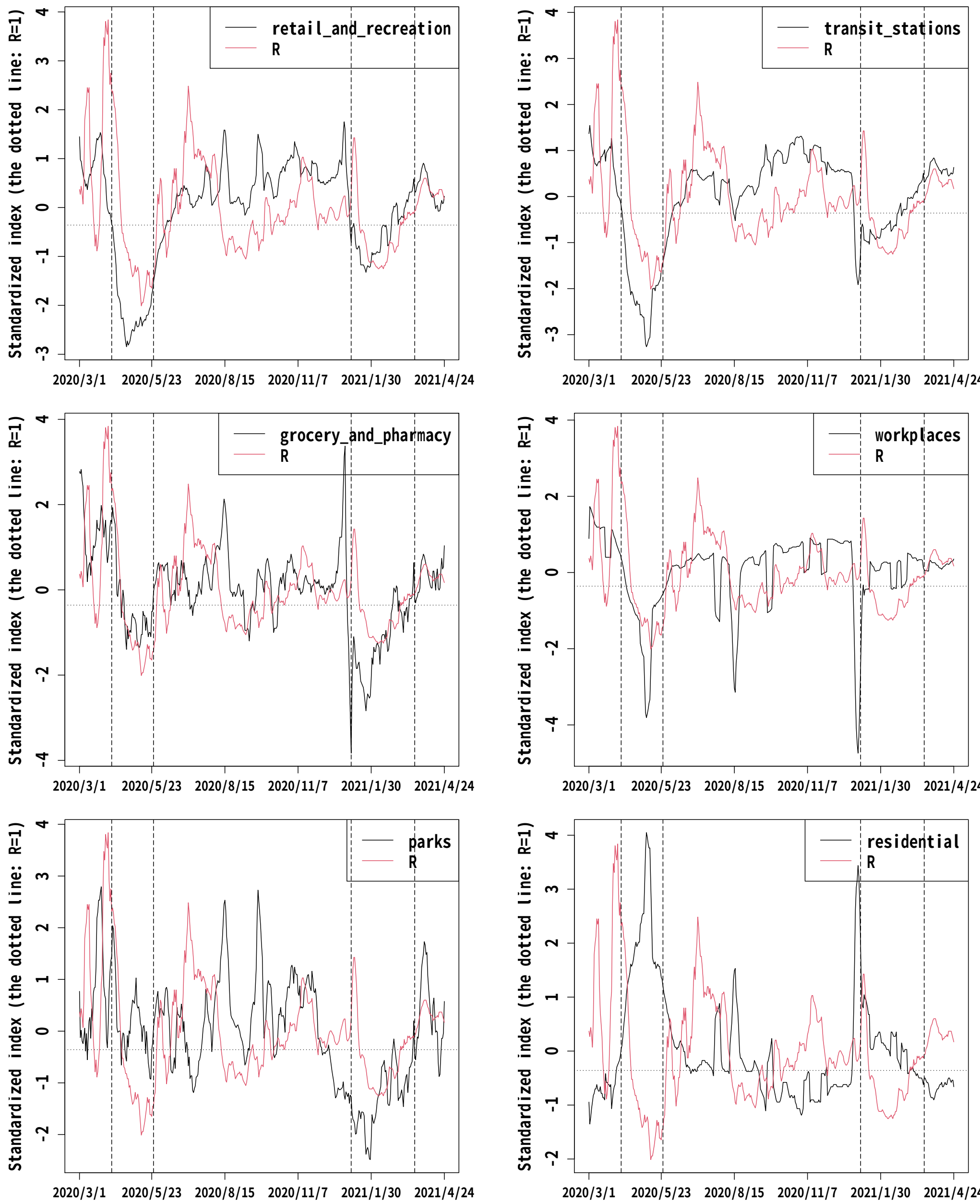
Notes: In each panel, "R" indicates the effective reproduction number, and nominal household expenditure per household is a 7-day backward moving average. Vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021. All figures are standardized by their means and standard deviations. The horizontal dotted line indicates the value of the standardized index for the effective reproduction number equal to 1 in each panel.

Figure 2: Cross correlation function between the effective reproduction number and the 7-day moving average of nominal household expenditure for clothing and footwear



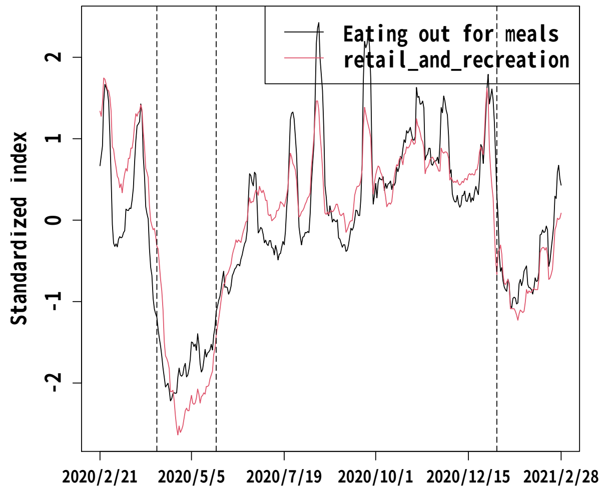
Notes: The figure shows the correlation coefficient between the contemporaneous effective reproduction number and lagged 7-day backward moving averages of nominal household expenditure per household for clothing and footwear. On the horizontal axis, negative lags are leads. Horizontal dashed lines are the 95% confidence interval for correlations between independent white noises. The sample period is from March 1, 2020, to February 28, 2021, as the effective reproduction number is available only from March 1, 2020.

Figure 3: Effective reproduction number and 7-day moving averages of mobility measures



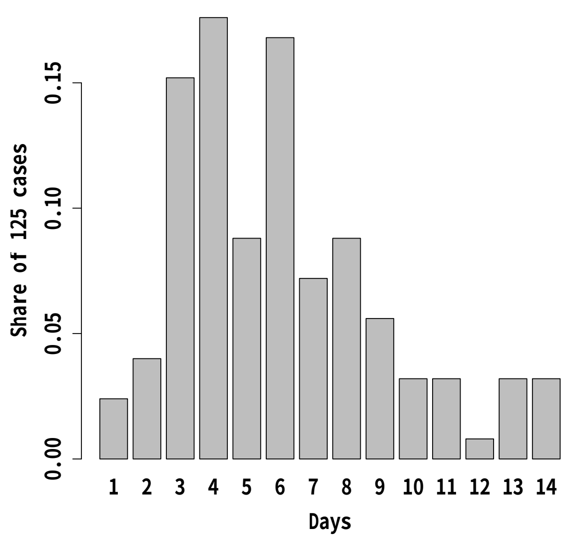
Notes: In each panel, "R" is the effective reproduction number, and the measure of mobility is a 7-day backward moving average. All figures are standardized by their means and standard deviations. Vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021.

Figure 4: 7-day moving averages of mobility in retail and recreation and real household expenditure per household on eating out for meals



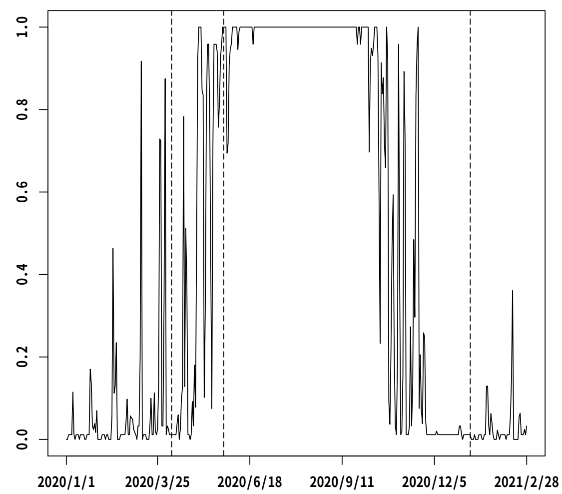
Notes: The figure plots retail_and_recreation in the COVID-19 Community Mobility Reports from Google and real household expenditure per household on eating out for meals. Both figures are 7-day backward moving averages, and standardized by their means and standard deviations. Vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021.

Figure 5: A sample distribution of incubation periods in Japan



Source: Sugishita, Kurita, Sugawara, and Ohkusa (2020).

Figure 6: Dummy variable for absolute humidity



Notes: The figure plots the daily value of $D_{AH,t}$. Vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021.

Figure 7: Causal diagram

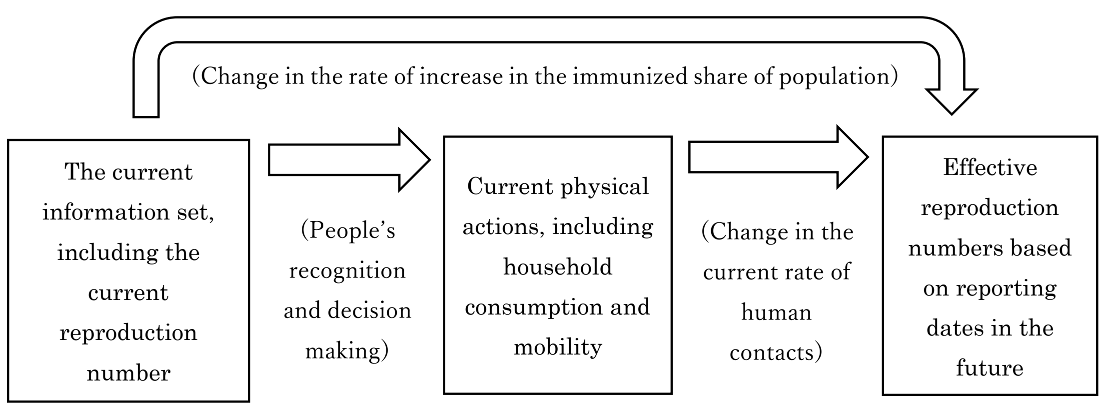
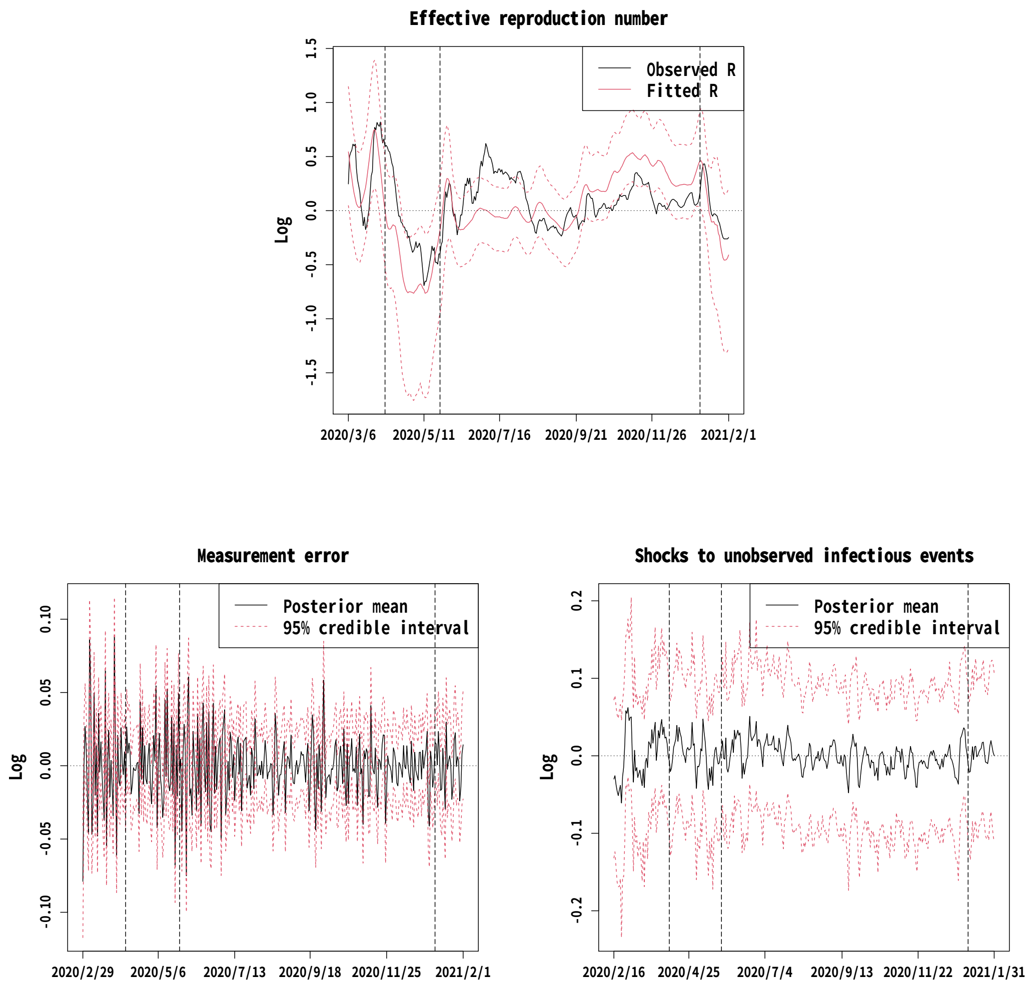
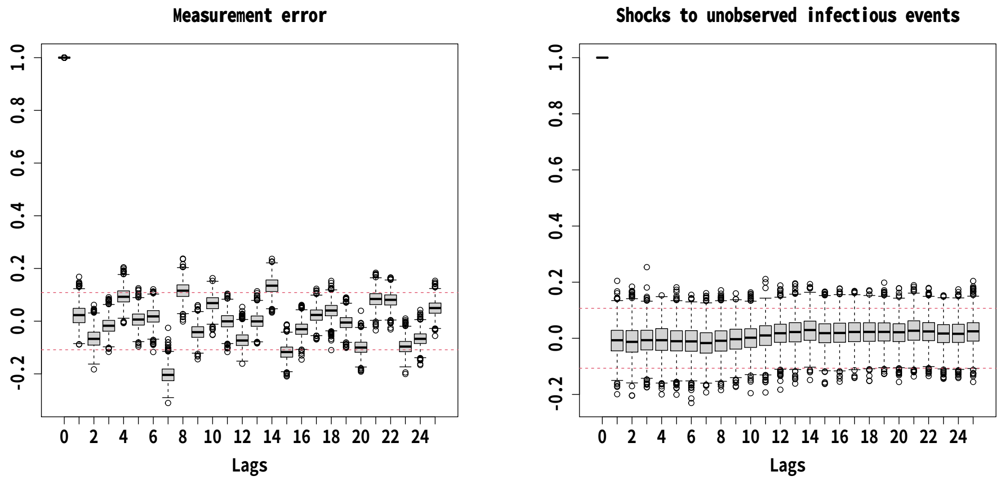


Figure 8: Fitted value of the effective reproduction number and residuals



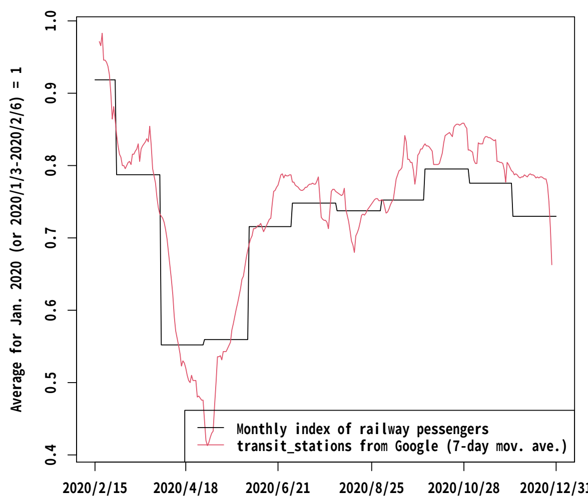
Notes: In the top panel, "Observed R" indicates the log of the observed effective reproduction number; and "Fitted R" indicates the fitted value of the log of the effective reproduction number in the regression model with 2020-21 data. Red dashed lines in each panel indicate the 95% credible interval. In the bottom panels, "Measurement error" and "Shocks to unobserved infectious events" indicate the values of η_t and ϵ_t, respectively. In both top and bottom panels, vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021.

Figure 9: Mcmc samples of auto-correlation functions of residuals



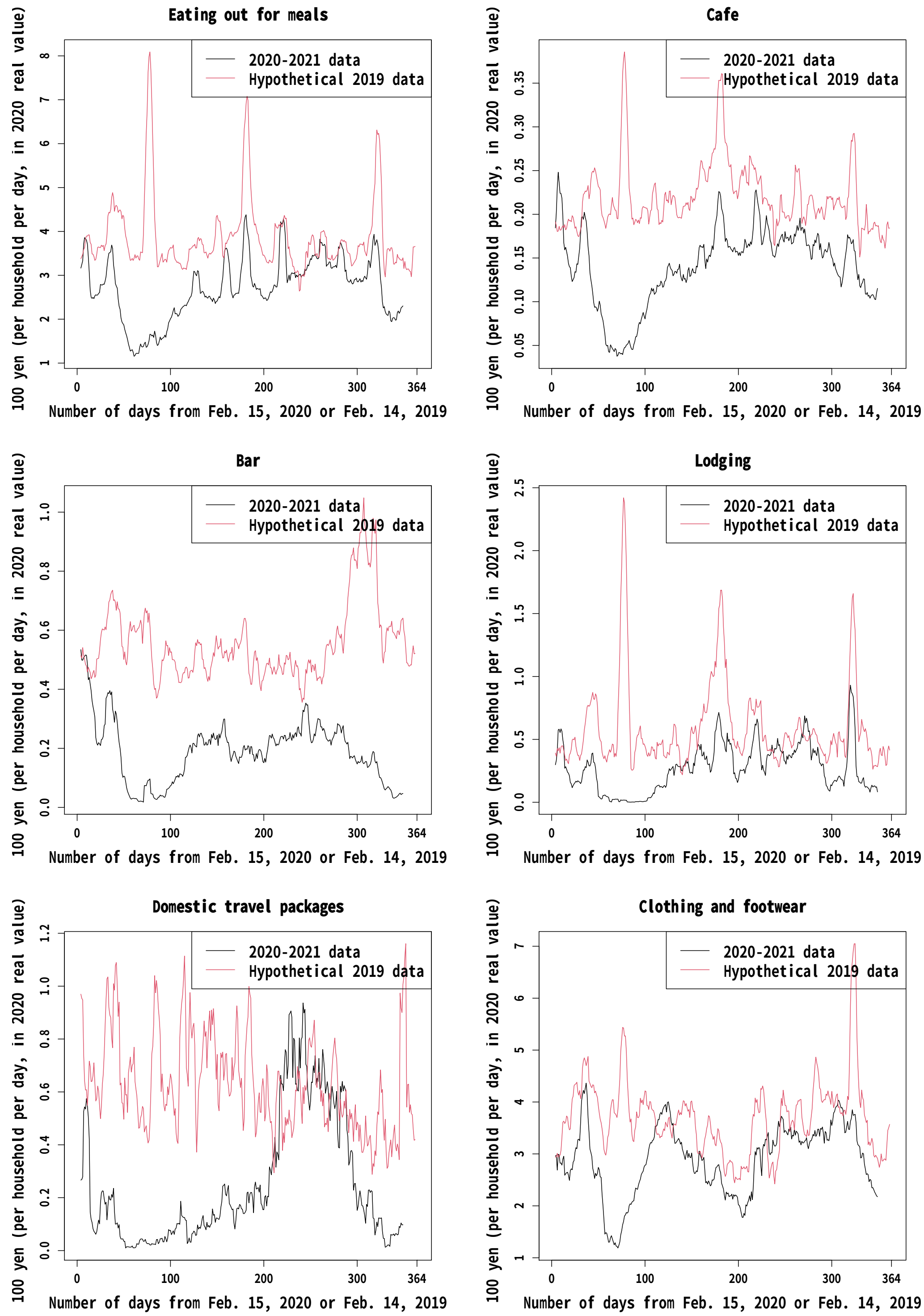
Notes: "Measurement error" and "Shocks to unobserved infectious events" indicate η_t and ϵ_t, respectively. For each lag, the grey box shows the range between 25% and 75% percentiles, and the black line in the middle of the box indicates the median. The whiskers extended above and below the box show the range between 25% percentile - 1.5*(75% percentile-25% percentile) and 75% percentile + 1.5*(75% percentile-25% percentile). Each circle indicates the value of an outlier outside this range.

Figure 10: The number of railway passengers and mobility in public transportation in 2020



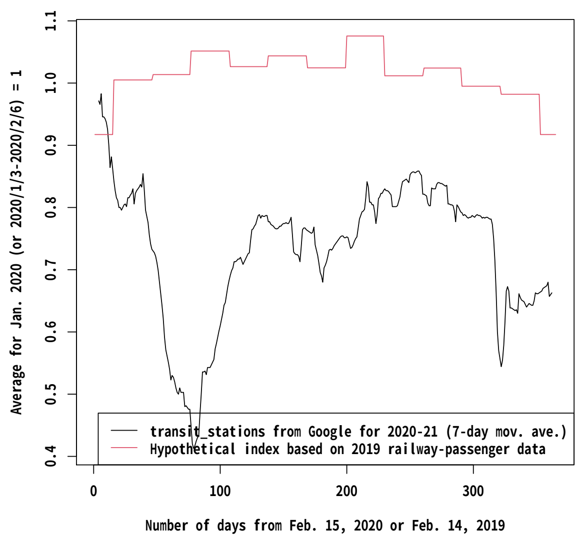
Notes: "transit_stations" is a measure of mobility in public transportation in the COVID-19 Community Mobility Reports from Google, which is available from February 15, 2020. For this measure, a 7-day centered moving average is shown in the figure. The index of railway passengers is constructed by dividing the monthly average of railway passengers in each month of 2020 by the monthly average in January 2020. The monthly value of this index is shown for each date within the same month.

Figure 11: Real household expenditures per household in 2019 and for 2020-21



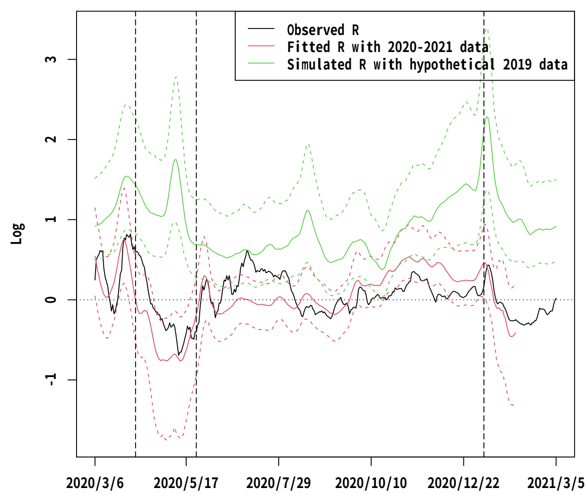
Notes: In each panel, the 2019 data start from February 14, 2019, and are connected with the data on January 1, 2019, after the year end, so that they loop for 365 days as hypothetical data without any restriction on household consumption or mobility. The 2020-21 data start from February 15, 2020, and end at January 31, 2021. All figures are 7-day centered moving averages.

Figure 12: Mobility in public transportation in 2019 and for 2020-21



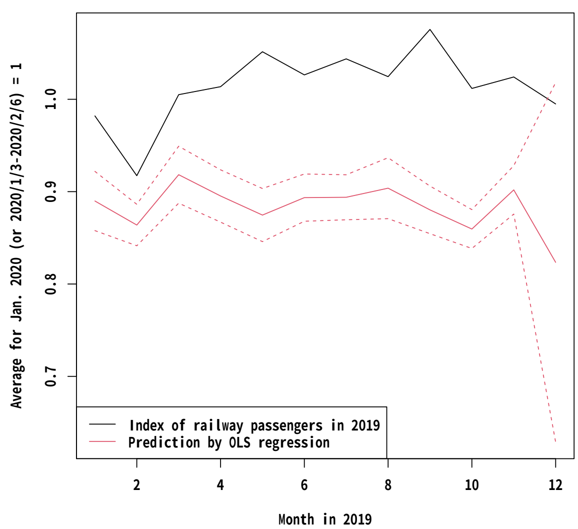
Notes: "transit_stations" is a measure of mobility in public transportation in the COVID-19 Community Mobility Reports from Google, which is available from February 15, 2020. The figure for this measure is a 7-day centered moving average. The index of railway passengers is constructed by dividing the monthly average of railway passengers in each month of 2019 by the monthly average in January 2020. The monthly value of this index is shown for each date within the same month. The index starts from February 14, 2019, and then is connected with its value on January 1, 2019, after the year end, so that it loops as a hypothetical index of mobility in public transportation without any restriction on household consumption or mobility.

Figure 13: Effective reproduction number without any restriction on household consumption or mobility



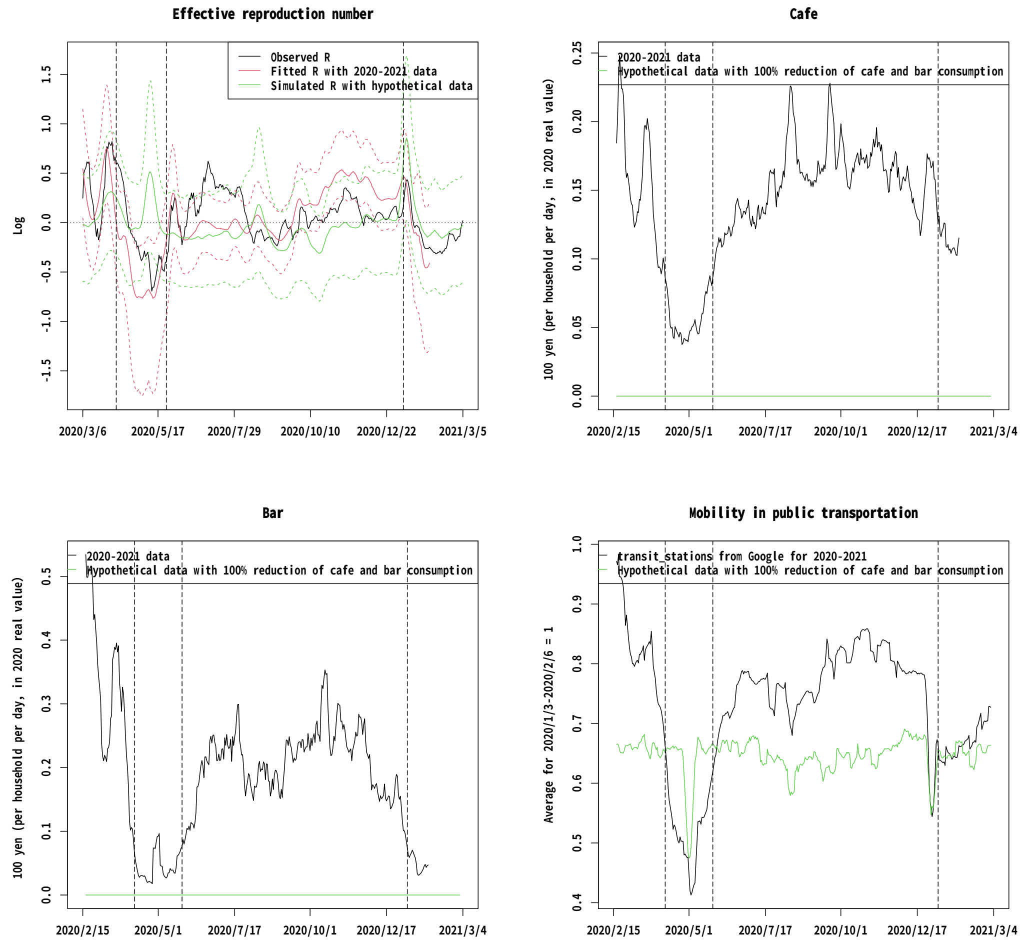
Notes: The vertical axis is the log of the effective reproduction number. "Observed R" is the log of the observed effective reproduction number. "Fitted R with 2020-2021 data" is the fitted value of the log of the effective reproduction number in the regression model with 2020-2021 data. "Simulated R with hypothetical 2019 data" is the daily value of $\ln R_t$ simulated by the regression model for the effective reproduction number with hypothetical values of explanatory variables based on 2019 data. Vertical dashed lines are the first and the last dates of two states of emergency: from April 7, 2020, to May 25, 2020; and from January 7, 2021, to March 21, 2021.

Figure 14: The observed and the fitted value of mobility in public transportation in 2019



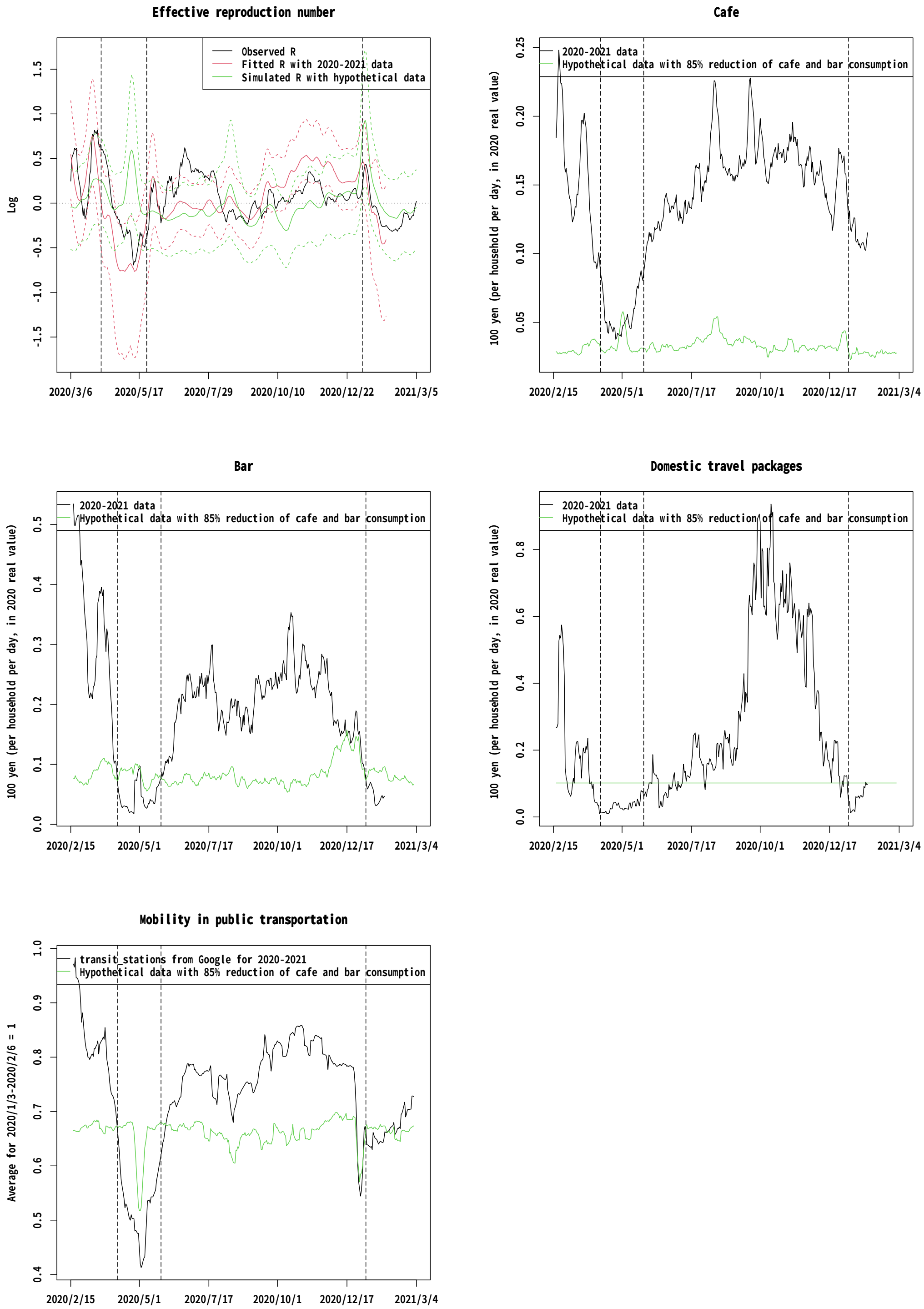
Notes: “Index of railway passengers in 2019” is the ratio of the monthly average of railway passengers in each month of 2019 to the January 2020 average. ”Prediction by OLS regression” is the fitted value generated by inserting 2019 data in the explanatory variables of the regression shown in Table 7, except that time dummies for the two states of emergency are set to zero. The daily fitted values are averaged out to compute the monthly average for each month in the figure. Dotted lines around ”Prediction by OLS regression” indicate the 95% confidence interval.

Figure 15: Effective reproduction number, restricted real household expenditures per household, and endogenized mobility in simulation 1



Notes: In each panel, “Observed R” is the log of the observed effective reproduction number; “Fitted R with 2020-2021 data” is the fitted values of the log of the effective reproduction number in the regression model with 2020-2021 data; “Simulated R with hypothetical data” is the simulated value of the log of the effective reproduction number in the regression model with “Hypothetical data with 100% reduction of cafe and bar consumption” in the other panels; “transit_stations from Google for 2020-2021” is a measure of mobility in public transformation in the COVID-19 Community Mobility Reports from Google from February 15, 2020; “Hypothetical data with 100% reduction of cafe and bar consumption” is an exogenously restricted real household expenditure item, or mobility in public transportation endogenized by the regression shown in Table 7, in simulation 1. Except the effective reproduction number, figures shown in each panel are 7-day centered moving averages.

Figure 16: Effective reproduction number, restricted real household expenditures per household, and endogeneized mobility in simulation 3



Note: See the notes for Figure 15.