Online Appendix
"COVID-19 Infection Spread and Human Mobility" by Shibamoto, Hayaki, and Ogisu

A.1 New positive cases of the COVID-19 infection

This subsection documents the historical time-series features of new positive COVID-19 infection cases in Japan. It especially focuses on non-stationarity and cyclicality in infection spread. Figure A1 shows a weekly time series of the confirmed new infection cases (top panel), log of new infection cases (middle panel), and log changes in infection cases (bottom panel) from the week of February 16, 2020, to the week of May 9, 2021. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan.[1]

The non-stationarity in the time series of the number of new infection cases is almost a certainty. The top panel of Figure A1 shows four peaks in the number of new infection cases, despite only three states of emergency. This shows that the expected value of the number of new infection cases tends to increase and decrease stochastically. That is, the time series of the number of new infection cases follows a non-stationary process.

However, the number of new infection cases seems to be cyclical. Figure A1 shows that the number of new infection cases decreased, despite no declaration of a state of emergency. Interestingly, the three peak-to-bottom periods are approximately the same (about a month and a half).[2] This similarity suggests that the decline in new infection cases tends to continue for approximately a month and a half, regardless of whether a state of emergency has been declared.

[1] In Japan, a state of emergency was declared three times during the sampling period: from April 7th to May 25th 2020, from January 8th to March 21st 2021, and from April 25th 2021 onward. The legal basis for the policy responses of the governors of the prefectures subject to the emergency measures, is the "Act on Special Measures for Pandemic Influenza and New Infectious Diseases Preparedness and Response." Under the first declaration of emergency, prefectural governors were able to request people to refrain from going out of their homes, and to request and instruct facility administrators of schools, social welfare facilities, and entertainment venues to restrict the use of those facilities, in accordance with the provisions of Article 45 of the Act. However, the Act does not stipulate any penalties for disobeying the instructions under Article 45, and Japan's curfew was extremely loose compared to the lockdowns in China, United States, and many European countries. On February 13th 2021, during the second declaration of the state of emergency, the Act has been amended to allow the prefectural governor to instruct the facility manager regarding restrictions. In addition, based on the newly established Article 79, facility managers who do not comply with the order will be subject to a fine of up to 300,000 yen.

[2] The first peak-to-bottom period is five weeks: from the week of April 12, 2020, to the week of May 17, 2020. The second peak-to-bottom period is seven weeks: from the week of August 2, 2020, to the week of September 20, 2020. The third peak-to-bottom period is seven weeks: from the week of January 10, 2021, to the week of February 28, 2021.

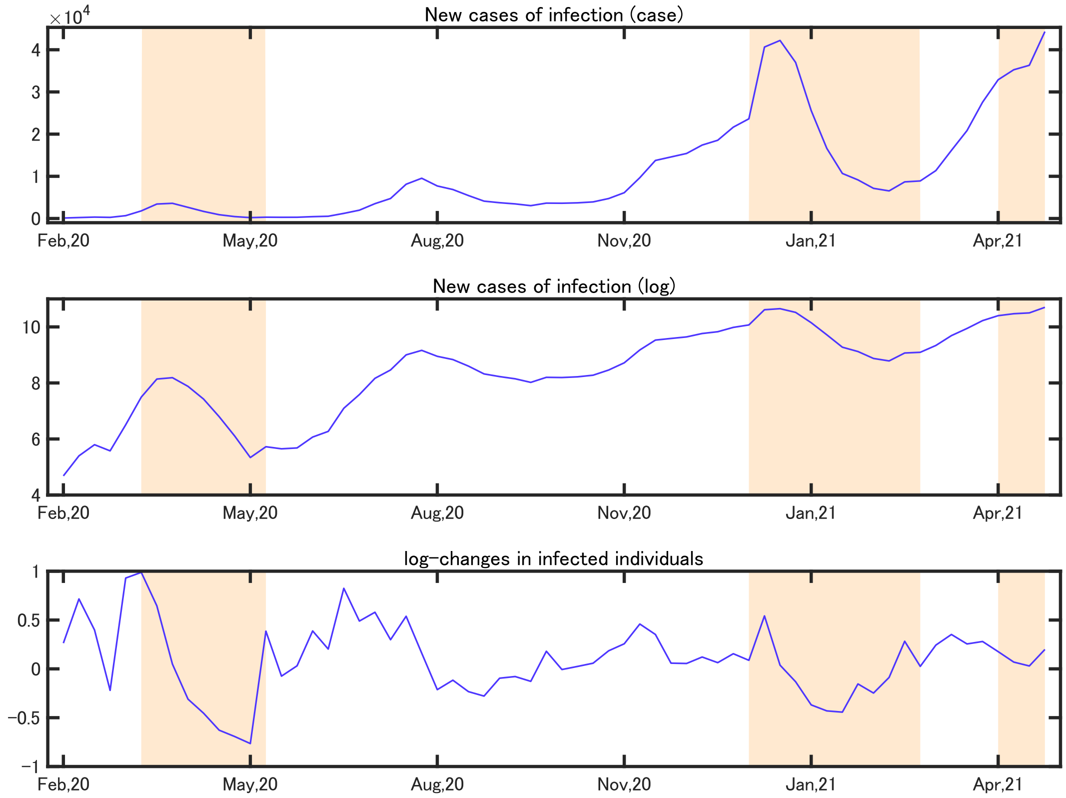

Figure A1: Weekly new positive cases of COVID-19 infection in Japan

Notes: The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period is the week of February 16, 2020, to the week of May 9, 2021.

Figure A2 shows a weekly time series of log changes in infection cases from the week of March 1, 2020, to the week of May 9, 2021 in Tokyo and Osaka.[3] The infection spread in each prefecture can be differentiated into spread at the same or different times. As seen in the figure, there appears to be a common trend in the infection cases across the two prefectures. We notice that two prefectures have experienced a sustained decrease in the infection cases during the declaration of the state of emergency, and a persistent increase in cases after the declaration was lifted. On the other hand, there are regional differences in the infection spread. The growth rate of new cases in the two prefectures can often be seen to temporarily diverge.

A.2 Human mobility index

We measure the movement of people using Google's mobility indices. The weekly mobility index we use matches the median of each week of the daily mobility index, to eliminate the effects

[3]Following Hoshi *et al.* (2021), we set the value of the log levels of new infection cases to -1 for observation with zero weekly cases

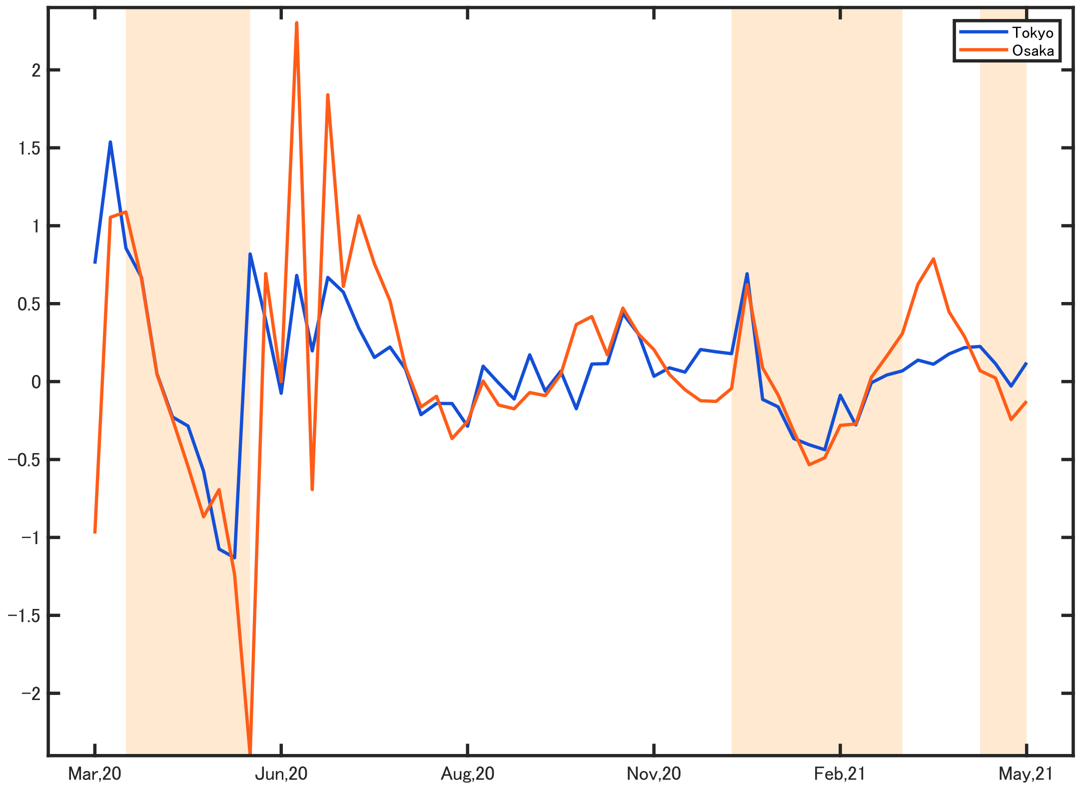

Figure A2: Weekly log changes in new positive cases of COVID-19 infection in Tokyo and Osaka

Notes: The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period spans from the week of March 1, 2020, to the week of May 9, 2021. Different colors represent different prefectures, where the navy blue is for Tokyo and the dark orange is for Osaka.

of holidays as much as possible. These mobility indices represent the rate of change in human mobility for six categories: retail & recreation, grocery & pharmacy, parks, transit stations, workplaces, and residential spaces.

First, we present the contemporaneous relationships among the mobility indices during the sample period. Table A1 reports the Pearson's correlation coefficients between two mobility indices. From Table A1, the six mobility indices are highly correlated. In particular, retail & recreation, transit stations, and workplaces are strongly intercorrelated, and residential shows a strong negative correlation. Hence, there are common components that help summarize the characteristics of these highly correlated variables.

To capture the common components among the mobility indices, we consider a model to summarize the time-series characteristics of these indices. Specifically, we assume that the representative measure for human mobility as the common factor of individual mobility indices, significantly drives the time-series patterns of the indices. Let $M_t = (m_{1t}, \cdots, m_{6t})'$ be a six-

Table A1: Correlation between the Google mobility indices

	[1]	[2]	[3]	[4]	[5]
[1] Retail & recreation					
[2] Grocery & pharmacy	0.528				
[3] Parks	0.292	0.516			
[4] Transit stations	0.857	0.293	0.200		
[5] Workplaces	0.394	-0.093	0.022	0.751	
[6] Residential	-0.749	-0.148	-0.108	-0.948	-0.854

Notes: This table shows the Pearson's correlation coefficient between the Google mobility indices from the week of February 16, 2020, to the week of May 9, 2021.

by-one vector of the mobility indices (retail & recreation, grocery & pharmacy, parks, transit stations, workplaces, residential). We consider the following factor model:

$$M_t = \lambda_y y_t + \lambda_1 + v_t, \tag{1}$$

where y_t is a single common factor, v_t is a six-by-one vector of mean-zero idiosyncratic components, and λ_y, λ_1 are six-by-one vectors of slope and intercept coefficients, respectively. We impose the first element of coefficient vector λ_y to be equal to one, to endow the factor y_t with interpretable units and signs, where one unit increase in y_t reflects an increase in the mobility index with retail & recreation by one.

We adopt the principal component (PC) approach to measure the representative measure for human mobility y_t. First, we calculate the first PC, pc_t, using the standardized (mean-zero and unit variance) six mobility indices. Next, we regress human mobility of retail & recreation m_{1t} on pc_t and the constant term:

$$m_{1,t} = \delta_{1,pc} pc_t + \delta_{1,1} + \nu_{1,t}. \tag{2}$$

where $\nu_{1,t}$ is an error term. We can then compute the representative measure for human mobility, satisfied with the restriction imposed in the factor model as $y_t = \hat{\delta}_{1,pc} pc_t$ where $\hat{\delta}_{1,pc}$ is an ordinary least squares (OLS) estimate of the parameter $\delta_{1,pc}$ in Equation (2). Since the representative measure for human mobility y_t is based on PC, it is a composite index of the mobility indices.

Table A2 shows the empirical results for the factor model (1) by OLS regression. From the table, the regression coefficients of four indices (retail & recreation, transit stations, workplaces, and [negative] residential) on the composite index are statistically significant at the 1% level. Moreover, the composite index explains most of the mobility index variations with the retail & recreation, transit stations, and residential categories, although the mobility indices with the grocery & pharmacy and parks categories are explained to a lesser degree, given the large share explained by the idiosyncratic components. Thus, the composite index as the representative measure for human mobility summarizes the common characteristics among mobility indices.

Table A2: Factor model for the Google mobility indices

Dependent variable:	Retail & recreation	Grocery & pharmacy	Parks	Transit stations	Workplaces	Residential
λ_y	1	0.17	0.47	1.34	0.92	-0.42
	(0.11)	(0.06)	(0.36)	(0.05)	(0.15)	(0.02)
λ_1	-0.138	0.001	-0.045	-0.254	-0.131	0.066
	(0.007)	(0.005)	(0.023)	(0.003)	(0.007)	(0.002)
R-squared	0.77	0.16	0.10	0.96	0.57	0.90

Notes: This table shows the results for the ordinary least squares regression (1) of the mobility index with each column category on the composite index of mobility and constant term. We obtain the composite index by scaling and signing the first principal component calculated using the Google mobility indices to the index for retail & recreation. The sample period is the week of February 16, 2020, through the week of May 9, 2021. The numbers in parentheses are Newey & West (1987) heteroskedasticity and autocorrelation robust standard errors for least squares with a four-week lag truncation.

Each panel in Figure A3 displays each mobility index time series and the composite index. We graphically confirm that the composite index tracks the common time-series patterns among the mobility indices, during the sample period.

Each panel of Figures A4 and A5 displays each mobility index time series and the composite index in Tokyo and Osaka, respectively.

A.3 Comparison of in-sample forecasting accuracy among a historical average, AR model, and VAR model for the log-changes in new infection cases

We assess multistep-ahead forecasts, computed by iterating forward the VAR model (10) in the main text. It examines current-, two-, four-, and eight-week forecast horizons. The H-week-ahead forecast is computed by $\sum_{h=0}^{H} \Delta\hat{\pi}_{t+h}$ where $\Delta\hat{\pi}_{t+h}$ is the h-week-ahead predictor of

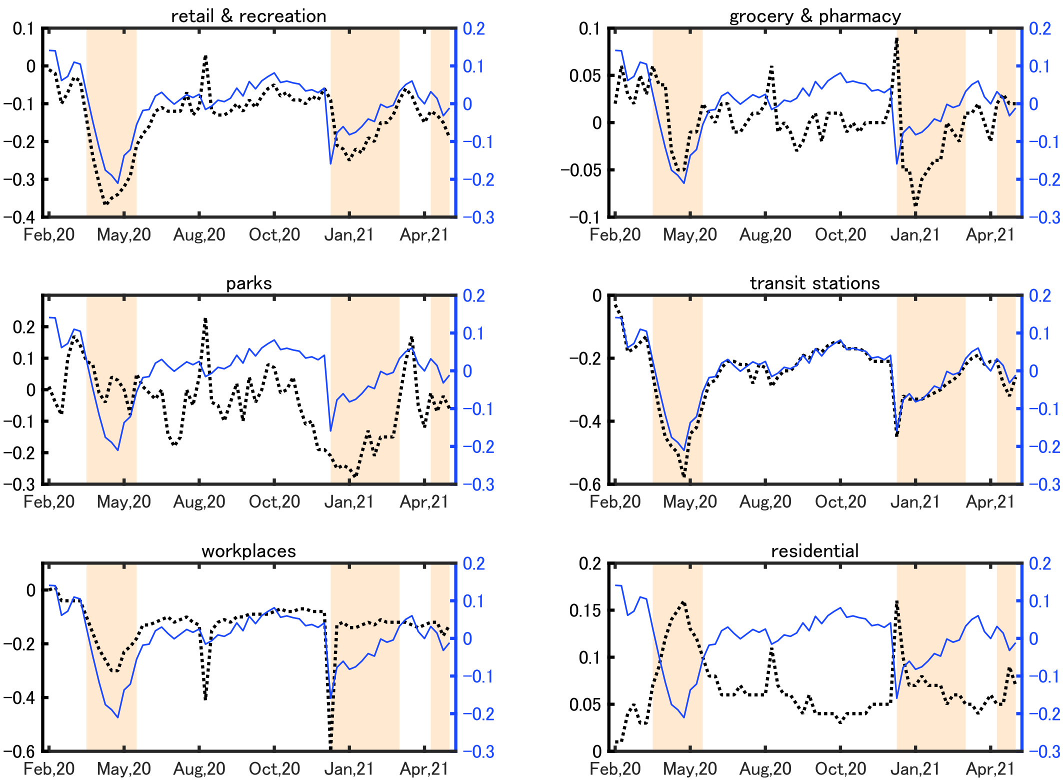

Figure A3: Weekly Google mobility index in Japan

Notes: The dotted line (left-hand scale) indicates the Google mobility index with each category. The solid line (right-hand scale) indicates the composite index of mobility. We obtain the composite index by scaling and signing the first principal component calculated using Google mobility indices to the index for retail & recreation. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period is the week of February 16, 2020, through the week of May 9, 2021.

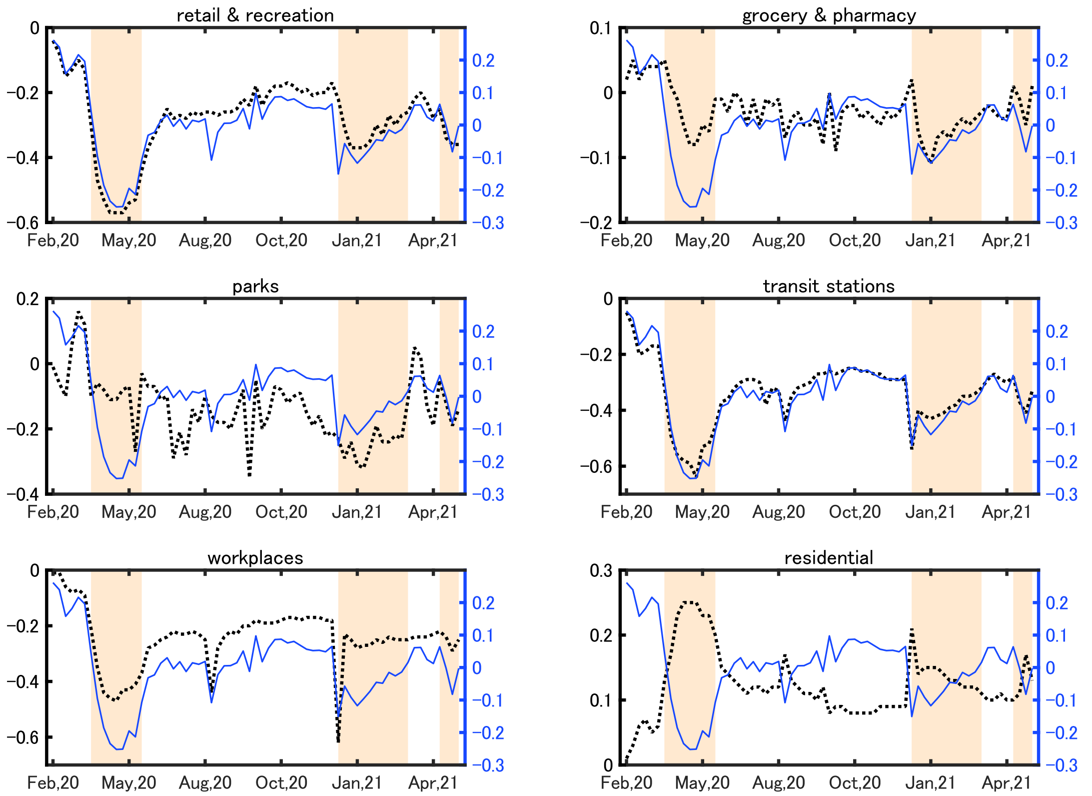

Figure A4: Weekly Google mobility index in Tokyo

Notes: The dotted line (left-hand scale) indicates the Google mobility index with each category. The solid line (right-hand scale) indicates the composite index of mobility. We obtain the composite index by scaling and signing the first principal component calculated using Google mobility indices to the index for retail & recreation. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period spans from the week of February 16, 2020, to the week of May 9, 2021.

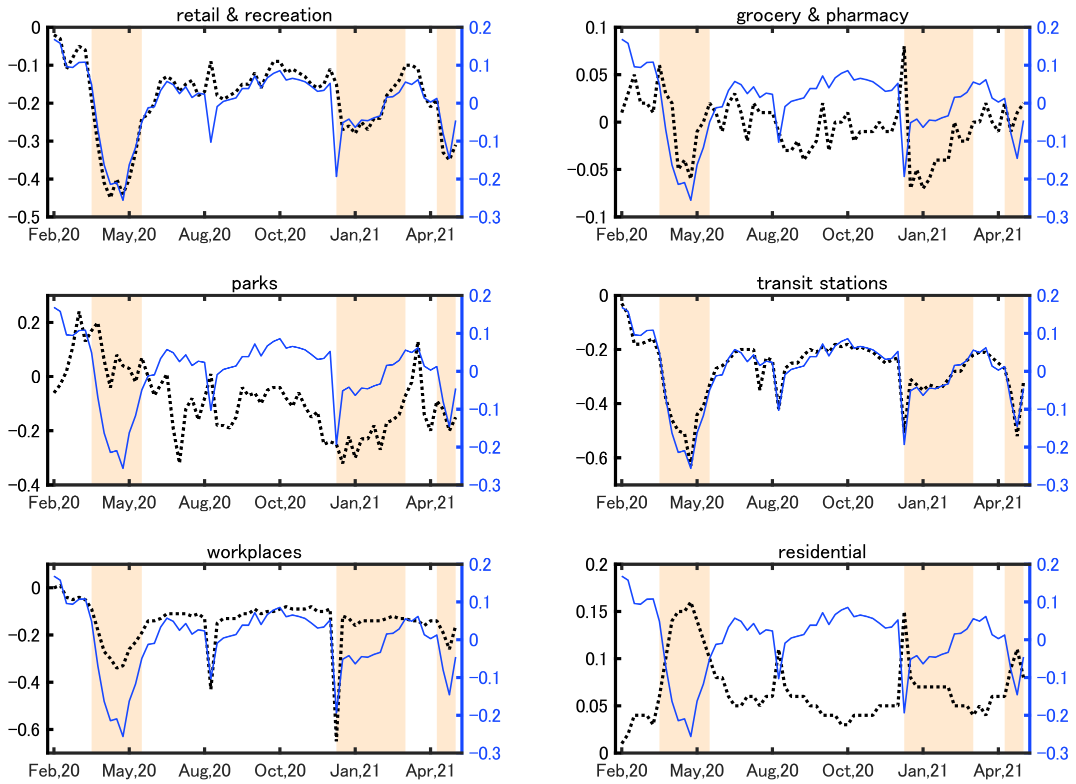

Figure A5: Weekly Google mobility index in Osaka

Notes: The dotted line (left-hand scale) indicates the Google mobility index with each category. The solid line (right-hand scale) indicates the composite index of mobility. We obtain the composite index by scaling and signing the first principal component calculated using Google mobility indices to the index for retail & recreation. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period spans from the week of February 16, 2020, to the week of May 9, 2021.

$\Delta\pi_{t+h}$ conditional on the information available at time $t-1$ using the forecasting models. As a comparison, we compute forecasts for a univariate AR with three-week lags and a historical average. Table A3 shows the mean square forecast error (MSE) and the percentage reduction in MSE for each of the forecasting models, relative to the historical average.

Accordingly, the predictions of the VAR model improve over the AR model and the historical average for any forecast horizon. Hence, the composite index of mobility contains valuable information for predicting the number of new infection cases.

The AR model forecast does not work adequately, especially for longer horizons. The upper four panels in Figure A6 shows the prediction values of the log changes in the new COVID-19 infection cases over the current (two, four, and eight) week(s) using the estimated parameters in the AR model. Although it has good predictive power in the shorter horizon, the predictive power of the AR model in the longer horizon is relatively poor.

Table A3: Prediction comparison for the log changes in infection cases

Horizon (H)	MSE_1	MSE_{AR}	R^2_{AR}	MSE_{VAR}	R^2_{VAR}
0	0.129	0.083	35.7	0.054	57.8
2	0.818	0.629	23.1	0.363	55.6
4	1.581	1.507	4.7	0.847	46.4
8	2.190	2.739	-25.1	2.015	8.0

Notes: This table shows the results from predicting the log changes in infection cases from the last week through the current (following two, four, and eight) week(s), conditioned on the information at the last week. MSE_1, MSE_{AR}, and MSE_{VAR} show the mean square forecast error (MSE) for the historical average, autoregressive (AR), and vector autoregressive model (VAR)forecasts, respectively. $R^2_{AR}(= 100 \times (1 - MSE_{AR}/MSE_1))$ and $R^2_{VAR}(= 100 \times (1 - MSE_{VAR}/MSE_1))$ measure the percentage reduction in MSE for the AR and VAR models, respectively, relative to the historical average. We set the lag length to three weeks in the AR and VAR estimation. Estimation samples in the AR and VAR model span the week of March 1, 2020, through the week of May 9, 2021. The forecasts are computed over the sample period depending on data availability of the actual H-week-ahead log changes in infection cases.

In contrast, the VAR model forecast works adequately over the sample period. The lower four panels in Figure A6 shows the prediction values of the log changes in the new COVID-19 infection cases over the forecast horizon using the estimated parameters in the VAR model. The predicted values explain the actual time-series pattern from March 2020 to May 2021, though the predictions after the state (exit) of (from) emergency declaration in early April (June) 2020 are slightly poor. This result suggests that the proposed VAR model simply and thriftily captures the dynamics of the log changes in new infection cases.

A.4 Estimation of the COVID-19 infection–mobility trade-off in Japan using each Google mobility index

As well as the benchmark, we estimate the following specification of the infection–mobility trade-off:

$$\Delta\pi_{t+1} = \kappa y_{t-1} + \iota + \epsilon_{\pi,t+1}. \tag{3}$$

For human mobility y_t, we use the Google mobility index with each category (retail & recreation, grocery & pharmacy, parks, transit stations, workplaces, residential) and the composite index of mobility. The sample period spans from the week of February 23rd, 2020, to the week of May 2nd, 2021.

Table A4 reports estimation results for the infection–mobility trade-off (3). The coefficients κ describing the response of the one-week-ahead log changes in infection cases, to the one-week lag

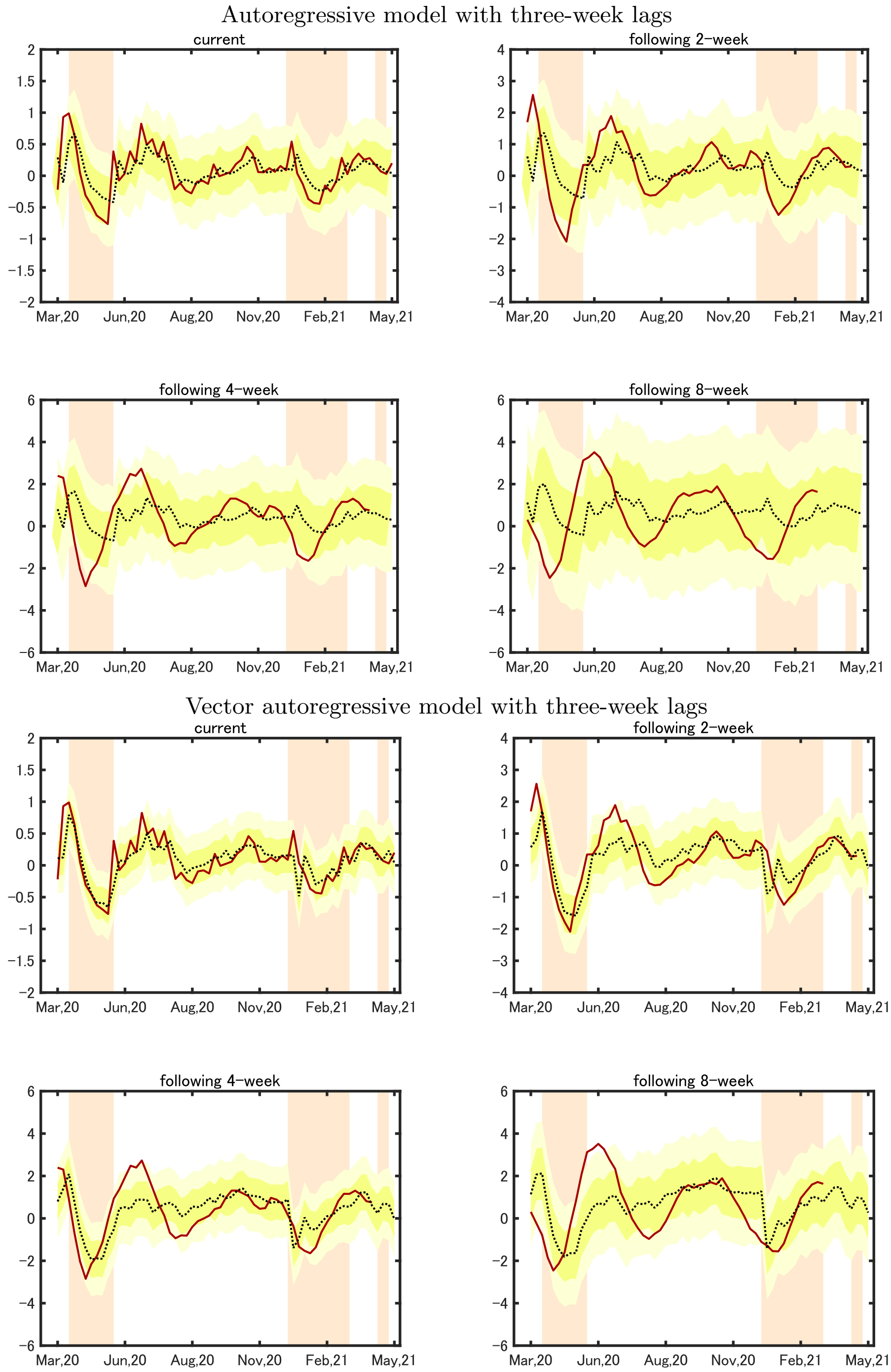

Figure A6: Prediction comparison for the log changes in infection cases

Notes: The solid line indicates the actual log changes in the new positive COVID-19 infection cases from the last week through the current (following two, four, or eight) week(s). The dotted line indicates the prediction values of the log changes, conditioned on the information at the last week; it is calculated using the parameter estimates of the autoregressive (AR) or vector autoregressive (VAR) model. We set the lag length to three weeks in the AR and VAR estimations. Dark and light yellow areas respectively denote ± 1 and ± 2 mean squared prediction error bands, calculated using 1,000 bootstrap samples. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. Estimation samples in the AR and VAR models span from the week of March 1, 2020, to the week of May 9, 2021.

Table A4: Estimation results for the COVID-19 infection–mobility trade-off in Japan

Dependent variable:	Retail & recreation	Grocery & pharmacy	Parks	Transit stations	Workplaces	Residential	Mobility CI
κ	2.67	4.97	0.76	2.33	1.76	-6.22	3.28
	(0.60)	(1.33)	(0.60)	(0.41)	(0.90)	(1.43)	(0.54)
ι	0.46	0.08	0.12	0.68	0.32	0.50	0.09
	(0.11)	(0.06)	(0.08)	(0.12)	(0.13)	(0.11)	(0.05)
Adj-R^2	0.38	0.18	0.04	0.40	0.18	0.30	0.43

Notes: This table shows the ordinary least squares regression results (3) of the log changes in infection cases on the mobility index with each column category and constant term. Mobility CI denotes the composite index of mobility. We obtain the composite index by scaling and signing the first principal component, calculated using Google mobility indices, to the index for retail & recreation. The sample period spans from the week of February 23, 2020, to the week of May 2, 2021. The numbers in parentheses are Newey & West (1987) heteroskedasticity and autocorrelation robust standard errors for least squares with a four-week lag truncation.

of human mobility, are positive and statistically significant for human mobility under the retail & recreation, grocery & pharmacy, transit stations, and workplaces categories. The coefficient is negative and statistically significant for human mobility under the residential category; the relationship is statistically insignificant for human mobility under the parks category.

Further, there are some differences in the explanatory power of the infection Phillips curve among the categories of the mobility index. In particular, the specification using the mobility index, especially with the parks and workplaces categories, has a lower R-squared than that using the mobility index with the retail & recreation, transit station, and residential categories. Thus, human mobility in parks and workplaces may be limited in explaining the time-series pattern of COVID-19 infection cases. These results are consistent with Nagata *et al.* (2021), where the mobility changes in nightlife and residential places are significantly associated with COVID-19 infection cases. Meanwhile, the association of mobility changes in workplaces is statistically insignificant. Moreover, the specification using the composite index of mobility has a higher R-squared than that using individual mobility indices. Hence, using the composite index improves the fit of the infection–mobility trade-off, by removing possible measurement errors for human mobility in individual mobility indices.

A.5 VAR-IV model

In this subsection, we describe the derivation of the VAR–IV model from a simultaneous equations system of infection-mobility trade-off and mobility demand.

We consider the following simultaneous equations system;

$$\Delta\pi_t = \zeta^{'} W_t + \xi_{a,t} + \phi\xi_{md,t}, \tag{4}$$

$$y_t = \gamma^{'} W_t + b\Delta\pi_t + \xi_{md,t}, \tag{5}$$

where W_t is a vector of control variables comprising the one- to three-week lags of y_t and $\Delta\pi_t$ and a constant. $\xi_{a,t}$ with a mean of zero and a variance of σ_a^2 represents an anxiety shock. $\xi_{md,t}$ with a mean of zero and a variance of σ_{md}^2 represents a mobility demand shock. $\xi_{a,t}$ and $\xi_{md,t}$ are assumed to be serially uncorrelated and independent each other. Equation (4) is a variant of the infection–mobility trade-off (Equation (15) in main manuscript) conditioning on the information at time $t-1$. $\xi_{a,t} + \phi\xi_{md,t}$ reflects the stochastic term of the new infection production, that is, unexpected changes in new infection cases at time t. Equation (5) is the mobility demand (Equation (16) in main manuscript).

We can express the structural VAR representation as follows;

$$A(L)X_t = a_0 + \Theta_a\xi_{a,t} + \Theta_{md}\xi_{md,t}, \tag{6}$$

where $\Theta_a = (\Theta_{1a}, \Theta_{2a})^{'} = (1, b)^{'}, \Theta_{md} = (\Theta_{1md}, \Theta_{2md})^{'} = (\phi, 1 + b\phi)^{'}$.

To identify the model, we use an *external* instrument from the VAR system that captures surprise variation in the new infection cases that people perceive as the change in infection risk. We can identify b under the premise that there is an instrument z_t for the anxiety shock that satisfies the following conditions:

$$E[\xi_{a,t} z_t] \neq 0, \text{ and} \tag{7}$$

$$E[\xi_{md,t} z_t] = 0. \tag{8}$$

Thus, we estimate the regression (5) by an IV estimation to obtain the estimate of b. Given the estimate of b, we can identify the mobility demand shocks as the residuals obtained from the IV regression. Thus, we estimate the regression (18) in the main manuscript to obtain the estimate of Θ_{md}.

A.6 The information content of the log level of weekly cases for the dynamics of human mobility in the VAR model

We consider a model referring to the specification by Chernozhukov *et al.* (2021) and Hoshi *et al.* (2021) that adds log change and log level of weekly cases at time t, $\Delta\pi_t$ and π_t, respectively, to the equation for the human mobility in the reduced-form VAR model;

$$y_t = b_1 \Delta\pi_t + b_2 \pi_t + \gamma^{'} W_t + \tilde{e}_{y,t}. \tag{9}$$

where W_t is a vector of control variables comprising the one- to three-week lags of y_t and $\Delta\pi_t$ and a constant, and $\tilde{e}_{y,t}$ is the error term orthogonal to $\Delta\pi_t$ and π_t. We estimate the regression (9) with and without $\Delta\pi_t$ and π_t as independent variables by an OLS, and compare the fit of the models.

Table A5 shows the empirical results. In the column (1), neither log changes nor log levels in infection cases at week t are included in the explanatory variables in Equation (9), i.e., the equation for the human mobility in the reduced-form VAR model. In column (2), the independent variable in the equation is the log changes in infection cases, i.e., the specification of the mobility demand in Equation (16) in the main manuscript. In column (3), the independent variable in the equation is the log levels in infection cases. In column (4), the independent variables in the equation are both the log changes and the log levels in infection cases. As shown in the column (1) of the Table, we find that about 70% of the variation in the composite index of mobility can be explained by the lag values of log changes in the new infection cases and the mobility index. As shown in the column (2) of the Table, the coefficient of the growth rate of cases at time t is statistically positive and significant, it does not have much additional explanatory power. As shown in the columns (3) and (4) of the Table, the coefficient of the log level of cases at time t is not statistically significant. This implies that adding the log level of weekly cases in our VAR model, plays a limited role in explaining the dynamics of the composite index of mobility.

Table A5: OLS Estimation results for the composit index of mobility with/without the log level of the new infection cases

Dependent variable:	Mobility CI			
	(1)	(2)	(3)	(4)
$\Delta\pi_t$	-	0.027	-	0.025
	-	(0.013)	-	(0.014)
π_t	-	-	-0.003	-0.002
	-	-	(0.004)	(0.004)
Adj-R^2	0.693	0.697	0.691	0.693

Notes: The dependent variable is the composite index of mobility. In column (1), both log changes and log levels in infection cases at week t are not included in the explanatory variables. In column (2), the independent variable is the log changes in infection cases. In column (3), the independent variable is the log levels in infection cases. In column (4), the independent variables are both the log changes and the log levels in infection cases. The constant and one- to three-week lags of log changes in infection cases, and the composite index of mobility are included as control variables in the linear regression model. We estimate the regression by ordinary least squares. The numbers in parentheses are White (1980) heteroskedasticity-robust standard errors. Coefficients and standard errors for the control variables are not reported. The sample period spans from the week of March 1, 2020, to the week of May 9, 2021.

A.7 The search volume of "感染者数" in Google

Figure A7 plots the weekly log changes in the search volume of the Japanese term "感染者数" (number of infected individuals *in English*) over time.[4]

A.8 Robustness check and sensitivity analysis

In this subsection, we re-estimate the model under several alternative settings to examine the robustness and sensitivity of our empirical analyses.[5]

Additional principal component for the Google mobility indices

We examine the robustness of the representative measure of human mobility. Although we assume that the time-series patterns of the Google mobility indices are significantly driven by their first PC, one may suspect a misspecification of our model for new infection cases and human behavior, due to the possibility that the other PCs also contain informational content about beneficial infection dynamics. Thus, we describe additional results that show that the first PC can sufficiently summarize the common features of the mobility indices using the factor

[4]We retrieved the search volume data from Google Trends (`https://trends.google.co.jp/trends/?geo=JP`) on June 1, 2021.

[5]Detailed results of the following exercises can be obtained from the corresponding author upon request.

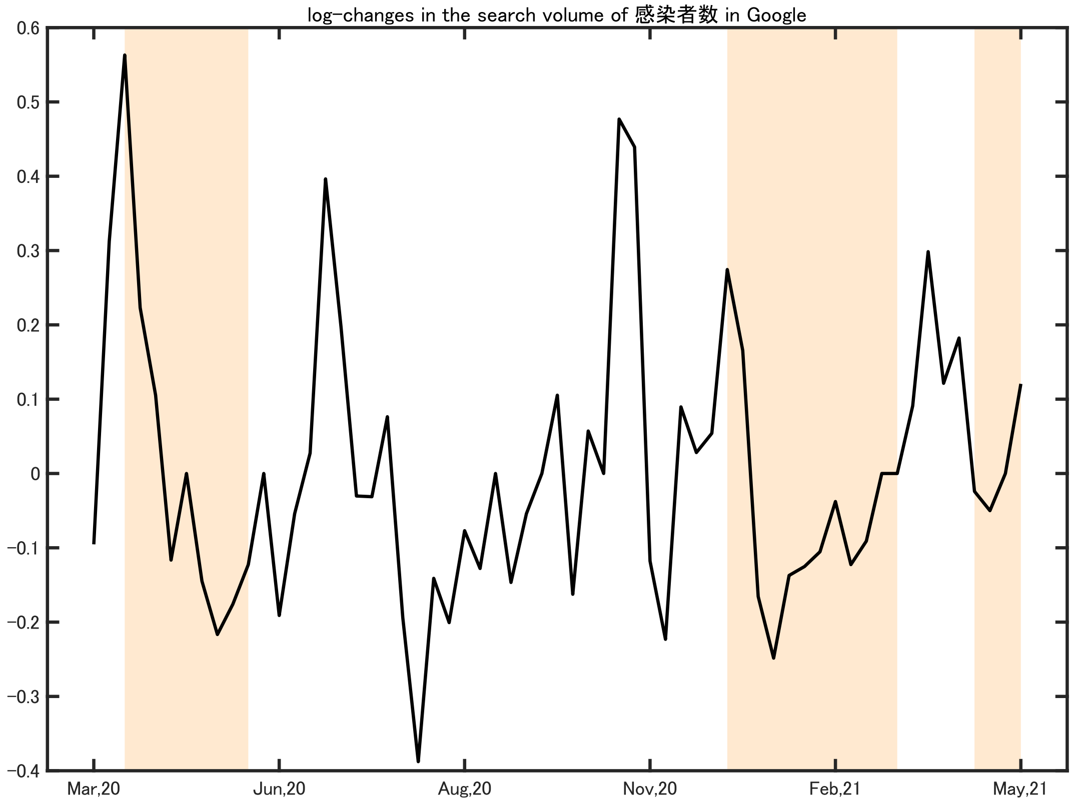

Figure A7: Weekly log changes in the search volume of “感染者数” (number of infected individuals *in English*) in Google

Notes: The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. The sample period spans from the week of March 1, 2020, to the week of May 9, 2021.

model, and the additional PC is limited in explaining the infection dynamics.

One way to justify the assumption that the number of common factors is equal to one is to check the information criteria for determining the number of factors. Specifically, we use the information criteria proposed by Ahn & Horenstein (2013), which suggests that the preferred model is the one that maximizes the information criteria, defined as follows:

$$IC(k) = \frac{\log(SSR(k-1))/\log(SSR(k))}{\log(SSR(k))/\log(SSR(k+1))}, \tag{10}$$

where $SSR(k)$ is the sum of squared residuals, $v^k_{i,t} = m_{i,t} - (\lambda^k_{i,y1}y_{1,t} + \cdots + \lambda^k_{i,yk}y_{k,t} + \lambda_{i,1})$ for $i = 1, \cdots, 6$ given the number of factors, k, such that

$$SSR(k) = \frac{1}{6T}\sum_{t=1}^{T}\sum_{i=1}^{6} v^{k2}_{i,t}. \tag{11}$$

Table A6: The number of common factors underlying the Google mobility indices

Number of factors (k)	0	1	2	3	4
$SSR(k)$	64.48	28.43	12.61	3.02	0.72
$IC(k)$	n.a.	0.94	0.58	-0.68	-11.97

Notes: $SSR(k)$ denotes the sum of squared residuals, $v_{i,t}^{k} = m_{i,t} - (\lambda_{i,y1}^{k} y_{1,t} + \cdots + \lambda_{i,yk}^{k} y_{k,t} + \lambda_{i,1})$ for $i = 1, \cdots, 6$ given the number of factors, k, defined as Equation (11). $IC(k)$ denotes the Ahn & Horenstein (2013) information criteria given the number of factors, k, defined as Equation (10). The sample period spans from the week of February 16, 2020, to the week of May 9, 2021.

Table A6 reports the calculated sum of squared residuals $SSR(k)$ and the Ahn & Horenstein (2013) information criteria $IC(k)$ for each number of factors k. The criteria suggest that we should adopt one common factor in summarizing the common features of the mobility indices using the factor model.

Second, although the inclusion of additional PCs increases explanatory power regarding mobility indices, it does not mean they are useful as a representative time-series measure for human mobility. Table A7 reports the empirical results for the factor model, including the two common factors estimated in the same manner as the benchmarks. In particular, from the table, the second PC strongly captures the characteristics of grocery & pharmacy and parks. Given that the mobility indices of these categories is limited in explaining the infection–mobility trade-off reported in Table A4, the second PC is unlikely to significantly capture the workings of the human mobility associated with the spread of infection. In fact, we extend the regression of the infection–mobility trade-off such that the second PC also affects the log changes in infection cases, confirming that the second PC has little explanatory power regarding the rate of new infection cases. Each panel in Figure A8, which displays the time series of each mobility index and the second composite index, shows that a large part of the fluctuations of the second PC is a significant spike, because of two long vacations: The former is the summer vacation in the week of August 9, 2020, and the latter is the winter vacation in the week of December 27, 2020. Although transitory changes in mobility during the long vacations may have affected new infection cases, such changes are not appropriate indicators to capture the time-series characteristics of the representative measure for human mobility, when modeling the intertemporal relationships between human mobility and new infection cases.

Next, we examine the information content of the second PC for the infection dynamics.

Table A7: Two factors model for the Google mobility indices

Dependent variable:	Retail & recreation	Grocery & pharmacy	Parks	Transit stations	Workplaces	Residential
λ_{y1}	1	0.17	0.47	1.34	0.92	-0.42
	(0.12)	(0.04)	(0.27)	(0.04)	(0.14)	(0.01)
λ_{y2}	1	1.32	4.10	-0.58	-2.32	0.47
	(0.50)	(0.14)	(0.78)	(0.18)	(0.58)	(0.04)
λ_1	-0.138	0.001	-0.045	-0.254	-0.131	0.066
	(0.008)	(0.002)	(0.013)	(0.002)	(0.008)	(0.001)
R-squared	0.82	0.81	0.63	0.97	0.83	0.98

Notes: This table shows the results for the ordinary least squares regression of the mobility index, with each column category on the composite indices of mobility constructed from the first and second principal components and constant term. We obtain the composite indices by scaling and signing the first and second principal components calculated using Google mobility indices to the index for retail & recreation. The sample period spans from the week of February 16, 2020, to the week of May 9, 2021. The numbers in parentheses are Newey & West (1987) heteroskedasticity and autocorrelation robust standard errors for least squares with a four-week lag truncation.

Table A8 shows the forecasting assessment of multistep-ahead forecasts, computed by iterating forward the three-variable VAR model comprising the log changes in the infection cases and the first and second composite indices for mobility, conducted in the same manner as the benchmarks. Four panels in Figure A9 show the prediction values of the log changes in the new infection cases, using the estimated parameters in the three-variable VAR model.

While the inclusion of the second composite index has contributed to some improvement in the model—specifically, reducing the prediction error—it has not significantly impacted the future movement of new infection cases. Comparing the results in Tables A3 and A8, the three-variable VAR model has a somewhat better prediction accuracy regarding in-sample prediction than the bivariate VAR model. However, comparing Figures A6 and A9, it is not apparent that the three-variable VAR model predictions have changed significantly from the bivariate VAR model predictions. Thus, although the additional PC contributed to reducing the prediction error of the VAR model, it does not change the dynamics of the VAR model. That is, the inclusion of the second composite index in the time-series model is not expected to be significant in analyzing the dynamic causal relationship between the number of infection cases and human mobility.

Finally, we conduct a sensitivity analysis, which includes the second composite index in our macroeconometric model. Specifically, we rerun the VAR model, including one- to three-week

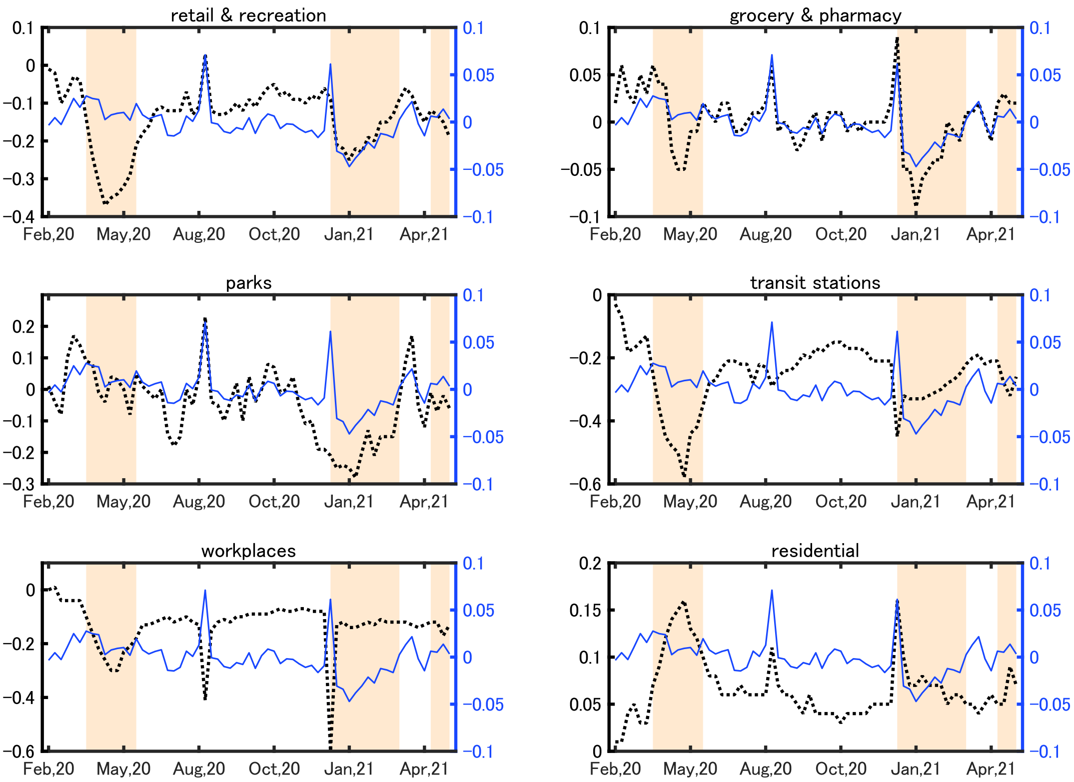

Figure A8: Weekly Google mobility index and the second composite index in Japan

Notes: The dotted line (left-hand scale) indicates the Google mobility index with each category. The solid line (right-hand scale) indicates the second composite index of mobility. We obtain the second composite index by scaling and signing the second principal component, calculated using Google mobility indices, to the index for retail & recreation. The orange shaded areas show the weeks coinciding with the period during which the state of emergency was declared in Japan. The sample period spans the week of February 16, 2020, to the week of May 9, 2021.

lags of the second composite index as exogenous variables. As expected, we confirm that the model produces, quantitatively and qualitatively, results that are similar to the benchmark results.

Assumption of epidemiological rigidity

We check the robustness for the empirical assessment of the infection–mobility trade-off based on the assumption of epidemiological rigidity. In particular, we assume epidemiological rigidity such that when infection increases due to increased mobility, it takes two weeks for the disease to develop under the benchmark specifications. One may attribute the sensitivity of the results to different settings for the rigidity because it may not take much time for the increased mobility to affect the spread of new infections. Thus, to consider the possibility that increased mobility

Table A8: Prediction comparison for the log changes in infection cases

Horizon	$MSE_{3varVAR}$	$R^2_{3varVAR}$
0	0.049	61.7
2	0.329	59.8
4	0.781	50.6
8	1.852	15.4

Notes: This table shows the results of predicting the log changes in infection cases from the last week through the current and the following two, four, and eight weeks conditioned on the information at the last week. $MSE_{3varVAR}$ shows the mean square forecast error (MSE) for the three-variable vector autoregressive (VAR) model forecast. $R^2_{3varVAR}(= 100 \times (1 - MSE_{3varVAR}/MSE_1))$ measures the percent reduction in the MSE for the three-variable VAR model relative to the historical average. We set the lag length to three weeks in the three-variable VAR estimation. Estimation samples in the three-variable VAR model span from the week of March 1, 2020, to the week of May 9, 2021. The forecasts are computed over the sample period, depending on data availability of the actual H-week-ahead log changes in infection cases.

may increase the new infection cases even after one week, we extend the regression of the infection–mobility trade-off, such that human mobility in period t also affects the log changes in infection cases during the period $t+1$:

$$\Delta\pi_{t+1} = \kappa_1 y_t + \kappa_2 y_{t-1} + \iota + \epsilon_{\pi,t+1}. \tag{12}$$

Along with the benchmark case in Equation (3), we use the Google mobility index with each category (retail & recreation, grocery & pharmacy, parks, transit stations, workplaces, and residential) and the composite index of mobility for human mobility y_t. Table A9 reports the estimation results for the *extended* infection–mobility trade-off in Equation (12).

Table A9 shows that our empirical assessment regarding the infection–mobility trade-off is most robust to the possibility that increased mobility may increase the new infection cases, after just one week. In particular, we confirm that the human mobility indices with transit stations and residential categories, and the composite index of mobility in week t, do not statistically contribute to the increase in the new infection cases in week $t+1$. This result suggests that our assumption of epidemiological rigidity in the benchmark model is realistic. We also confirm that the mobility in parks and workplaces is limited in explaining the time-series pattern of COVID-19 infection cases. Moreover, the specification using the composite index of mobility has a higher R-squared than that using individual mobility indices.

Note that there is a slight difference in the benchmarks. Regarding the retail & recreation

Three-variable vector autoregressive model with three-week lags

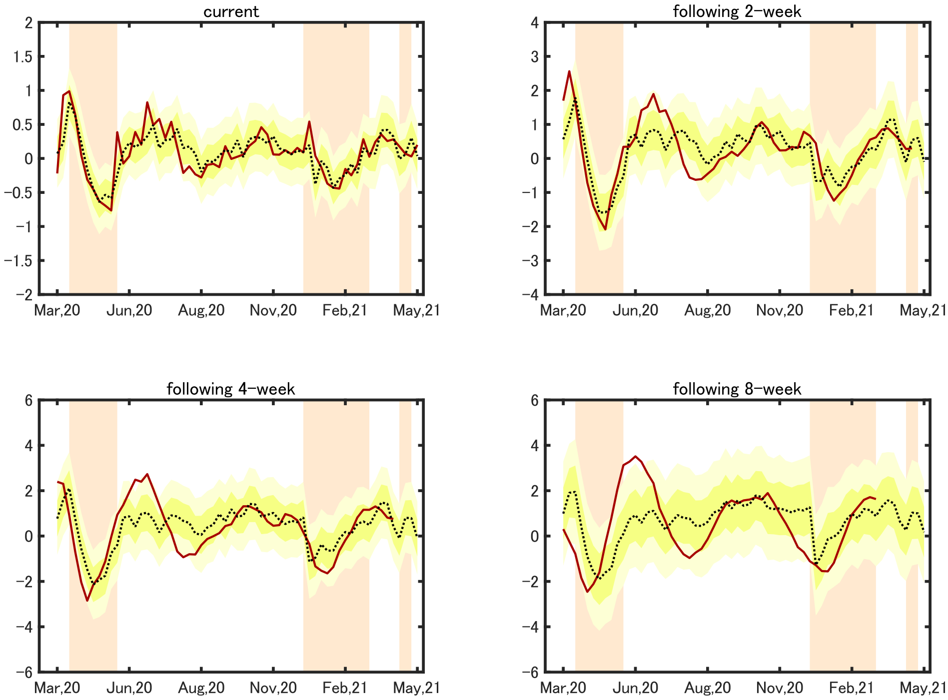

Figure A9: Prediction of the log changes in infection cases using a three-variable vector autoregressive model

Notes: The solid line indicates the actual log changes in new positive COVID-19 infection cases from the last week to the current (following two, four, and eight) week(s). The dotted line indicates the prediction values of the log changes, conditioned on the information at the last week. It is calculated using the parameter estimates of the three-variable vector autoregressive (VAR) model. We set the lag length to three weeks in the VAR estimation. Dark and light yellow areas denote ± 1 and ± 2 mean squared prediction error bands, calculated using 1,000 bootstrap samples, respectively. The orange shaded areas show the weeks coinciding with the period during the state of emergency declaration in Japan. Estimation samples in the three-variable VAR model span from the week of March 1, 2020, to the week of May 9, 2021.

and grocery & pharmacy categories, the mobility in week t contributes to a statistically significant increase in new infection cases in week $t+1$. In particular, the explanatory power of the model on the new infection cases, using the mobility index with the grocery & pharmacy category, has increased to a reasonable degree. Although we cannot explain why the explanatory power of the new infection rate increases only for a certain human mobility category, human mobility in the grocery & pharmacy category can explain the time-series pattern of the new infection cases.

Table A9: Estimation results for the *extended* COVID-19 infection–mobility trade-off in Japan

Dependent variable:	Retail & recreation	Grocery & pharmacy	Parks	Transit stations	Workplaces	Residential	Mobility CI
κ_1	1.36	4.86	0.15	0.25	0.63	-0.79	0.53
	(0.55)	(1.09)	(0.39)	(0.59)	(0.72)	(1.74)	(0.89)
κ_2	1.54	2.40	0.65	2.13	1.57	-5.64	2.86
	(0.79)	(1.19)	(0.51)	(0.48)	(0.63)	(1.15)	(0.63)
ι	0.49	0.09	0.12	0.70	0.38	0.51	0.09
	(0.11)	(0.05)	(0.08)	(0.14)	(0.16)	(0.13)	(0.05)
Adj-R^2	0.39	0.30	0.02	0.39	0.19	0.29	0.42

Notes: This table shows the results for the ordinary least squares regression (12) of the log changes in infection cases on the mobility index, with each column category and constant term. Mobility CI denotes the composite index of mobility. We obtain the composite index by scaling and signing the first principal component, calculated using Google mobility indices, to the index for retail & recreation. The sample period is the week of February 23, 2020, through the week of May 2, 2021. The numbers in parentheses are Newey & West (1987) heteroskedasticity and autocorrelation robust standard errors for least squares with a four-week lag truncation.

VAR settings

First, we check the sensitivity of the results to the lag length of the VAR model. In applications using the VAR model with long-run restrictions, one may attribute the sensitivity of the results to minor specification changes that are expected if identification is weak. In particular, when long-run restrictions are used to identify the VAR model, the lag length of VAR plays a role in identifying restrictions.[6] Even when we use the lag length of the VAR model from one week to five weeks, the models generate, qualitatively and quantitatively, results that are similar to those of a benchmark specification.

Second, we perform exercises incorporating measures for the strength of the government's reaction to the pandemic into the system. Specifically, we rerun the VAR model, including the current and one- to three-week lags of log changes of Oxford's Stringency Index, as exogenous variables. Surprisingly, none of the exercises changed the qualitative nature of the results.

Third, we perform exercises to incorporate some event dummies into the system. Specifically, we construct two long vacation dummies. The summer vacation dummy is set to one at the week of August 9, 2020, and zero elsewhere. The winter vacation dummy is set to one at the week of December 27, 2020, and zero elsewhere. It may not be appropriate to include the winter vacation

[6]See Faust & Leeper (1997) for details.

dummy as a control variable, due to the possibility of losing useful information regarding the identification of structural shocks pertaining to the Japanese people, who seemed to change their mobility patterns during the winter vacation at the end of 2020 (The following week, the Japanese government declared a second state of emergency, on January 7, 2021.). Nevertheless, the model produced quantitatively similar results for impulse response functions. However, the historical decomposition due to structural shocks around January 2021, was different from the benchmark, because the significant drop in the permanent shock at the week of December 27 disappeared in the alternative model.

References

Ahn, Seung C., & Horenstein, Alex R. 2013. Eigenvalue Ratio Test for the Number of Factors. *Econometrica*, **81**(3), 1203–1227.

Chernozhukov, Victor, Kasahara, Hiroyuki, & Schrimpf, Paul. 2021. Causal impact of masks, policies, behavior on early covid-19 pandemic in the U.S. *Journal of Econometrics*, **220**(1), 23–62.

Faust, Jon, & Leeper, Eric M. 1997. When Do Long-Run Identifying Restrictions Give Reliable Results? *Journal of Business and Economic Statistics*, **15**(3), 345–353.

Hoshi, Kisho, Kasahara, Hiroyuki, Makioka, Ryo, Suzuki, Michio, & Tanaka, Satoshi. 2021. Trade-off between job losses and the spread of COVID-19 in Japan. *Japanese Economic Review*, **72**(4), 683–716.

Nagata, Shohei, Nakaya, Tomoki, Adachi, Yu, Inamori, Toru, Nakamura, Kazuto, Arima, Dai, & Nishiura, Hiroshi. 2021. Mobility Change and COVID-19 in Japan: Mobile Data Analysis of Locations of Infection. *Journal of Epidemiology*, **31**(6), 387–391.

Newey, Whitney K., & West, Kenneth D. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**(3), 703–708.

White, Halbert. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, **48**(4), 817–38.