

**Supporting Information for**  
**Spotting epidemic keystones by  $R_0$  sensitivity analysis: High-risk**  
**stations in the Tokyo metropolitan area**

Kenta Yashima<sup>1,2</sup>, Akira Sasaki<sup>1,3</sup>

<sup>1</sup> Department of Evolutionary Studies of Biosystems, the Graduate University for Advanced Studies (SOKENDAI), Hayama, Kanagawa, Japan.

<sup>2</sup>Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, Nakano, Tokyo, Japan.

<sup>3</sup>Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria.

**Supporting Information S1. Derivation of the next generation matrix  $L$**

Derivation of the next generation matrix  $L$  for an infectious disease in a metropolitan area is given. By applying a linear approximation  $S_i^R(t) \cong N_i^R$  and  $S_{ij}^C(t) \cong N_{ij}^C$  to the integral forms of Eqs. (2), (5), we have the following renewal equations.

$$I_i^R(t) = \beta \int_0^t e^{-\gamma\tau} [2I_i^R(t-\tau) + I_i^H(t-\tau)] N_i^R d\tau \quad (\text{A1})$$

$$I_{ij}^C(t) = \beta \int_0^t e^{-\gamma\tau} [I_i^R(t-\tau) + I_i^H(t-\tau) + I_j^W(t-\tau)] N_{ij}^C d\tau \quad (\text{A2})$$

Here,  $I_i^H(t) \equiv \sum_j I_{ij}^C(t)$  and  $I_j^W(t) \equiv \sum_i I_{ij}^C(t)$  denote the number of infectious commuters in  $i$ -th home population and in  $j$ -th work population, respectively. By summing over indices  $i$  and  $j$  of Eq. (A2) and introducing a vector notation for non-commuting resident population ( $\mathbf{I}^R(t) \equiv (I_1^R(t), I_2^R(t), \dots, I_M^R(t))^t$ ), commuting home population ( $\mathbf{I}^H(t) \equiv (I_1^H(t), I_2^H(t), \dots, I_M^H(t))^t$ ), and commuting work population ( $\mathbf{I}^W(t) \equiv (I_1^W(t), I_2^W(t), \dots, I_M^W(t))^t$ ), Eqs. (A1), (A2) can be summarized in the  $3 \times 3$  block matrix form as

$$\begin{bmatrix} \mathbf{I}^R(t) \\ \mathbf{I}^H(t) \\ \mathbf{I}^W(t) \end{bmatrix} = \beta \int_0^t e^{-\gamma\tau} \begin{bmatrix} \mathbf{T}_{RR} & \mathbf{T}_{RH} & 0 \\ \mathbf{T}_{HR} & \mathbf{T}_{HH} & \mathbf{T}_{HW} \\ \mathbf{T}_{WR} & \mathbf{T}_{WH} & \mathbf{T}_{WW} \end{bmatrix} \begin{bmatrix} \mathbf{I}^R(t-\tau) \\ \mathbf{I}^H(t-\tau) \\ \mathbf{I}^W(t-\tau) \end{bmatrix} d\tau. \quad (\text{A3})$$

Here, element of this block matrix  $\mathbf{T}_{mn}$  is a  $M \times M$  matrix and denotes the transmission from type  $n$  population to type  $m$  population ( $m, n \in \{R, H, W\}$ ,  $R$ : non-commuting resident population,  $H$ : commuting home population,  $W$ : commuting work population), where each element of the matrices are given as  $[\mathbf{T}_{RR}]_{ik} \equiv 2N_i^R \delta_{ik}$ ,  $[\mathbf{T}_{RH}]_{ik} \equiv N_i^R \delta_{ik}$ ,  $[\mathbf{T}_{HR}]_{ik} \equiv N_i^H \delta_{ik}$ ,  $[\mathbf{T}_{HH}]_{ik} \equiv N_i^H \delta_{ik}$ ,  $[\mathbf{T}_{HW}]_{ik} \equiv N_{ik}$ ,  $[\mathbf{T}_{WR}]_{ik} \equiv N_{ki}$ ,  $[\mathbf{T}_{WH}]_{ik} \equiv N_{ki}$ , and  $[\mathbf{T}_{WW}]_{ik} \equiv N_i^W \delta_{ik}$  ( $\mathbf{0}$ :  $M \times M$  zero matrix). Here, it should be noted that the matrix  $\mathbf{T}_{mn}$  includes the information about the host population structure only and the epidemiological information is not included. Since  $\beta \int_0^t e^{-\gamma\tau} \mathbf{T}_{mn} \mathbf{I}^n(t-\tau) d\tau$  gives the number of infectious individual from type  $n$  population to type  $m$  population, the asymptotic ratio between these population types at exponential growth phase are given as  $\beta \int_0^\infty e^{-\gamma\tau} \mathbf{T}_{mn} d\tau = \beta/\gamma \mathbf{T}_{mn}$ . Given this the next generation matrix  $\mathbf{L}$  is given as a  $3 \times 3$  block matrix form as

$$\begin{aligned} \mathbf{L} &= \beta \int_0^\infty e^{-\gamma\tau} \begin{bmatrix} \mathbf{T}_{RR} & \mathbf{T}_{RH} & 0 \\ \mathbf{T}_{HR} & \mathbf{T}_{HH} & \mathbf{T}_{HW} \\ \mathbf{T}_{WR} & \mathbf{T}_{WH} & \mathbf{T}_{WW} \end{bmatrix} d\tau \\ &= \frac{\beta}{\gamma} \begin{bmatrix} \mathbf{T}_{RR} & \mathbf{T}_{RH} & 0 \\ \mathbf{T}_{HR} & \mathbf{T}_{HH} & \mathbf{T}_{HW} \\ \mathbf{T}_{WR} & \mathbf{T}_{WH} & \mathbf{T}_{WW} \end{bmatrix}. \end{aligned} \quad (\text{A4})$$

Here, each element of the block matrix gives the asymptotic ratio between different population types. Accordingly, the dominant eigenvalue of this next generation matrix  $\mathbf{L}$  gives the basic reproductive ratio  $R_0 = \rho(\mathbf{L})$  ( $\rho(\dots)$ : spectral radius).

For the dominant eigenvalue  $R_0$ , the elements of corresponding left and right eigenvector give the reproductive value and the relative ratio of exponentially growing infected population, respectively [1–3]. The reproductive value (i.e., element of left eigenvector  $(v_i^R, v_i^H, v_i^W)$ ) of each local population is given in Figure S2-A as a function of its local population size  $(N_i^R, N_i^H, N_i^W)$ . The relative ratio of exponentially growing infected population (i.e., element of right eigenvector  $(w_i^R, w_i^H, w_i^W)$ ) of each local population is given in Figure S2-B as a function of its local population size

$(N_i^R, N_i^H, N_i^W)$ . Both of the values increase as the local population size increase. This means that, a local population with a larger population size has a larger impact on the overall epidemic dynamics and also has a higher risk of infection. Furthermore, the most notable point is that, the results can be clustered into two distinct groups. This separation can be explained in the relation to the largest work population (i.e., working/studying area of Shinjuku station). For the results of non-commuting resident population and commuting home population (Figure S2-A2, 3, S2-B2, 3), the local populations in the upper cluster has at least one commuter who is working/studying at the largest work population. The horizontally layered colored structure can be clearly explained by the number of commuters to the largest work population. On the other hand, no one from the local populations in the lower cluster is working/studying at the largest work population. For the results of commuting population at work population (Figure S2-A1, S2-B1), the upper cluster is the largest work population itself and the lower cluster is consisted from other local populations. The clear distinction observed in the change in the basic reproductive ratio  $\delta R_0$  (see Figure 3), can be attributed to this distinction in the eigenvectors.

### Supporting Information S2. Final size of epidemic

Overall damage can be evaluated by the global final size of epidemic  $\Psi$ , which is defined as a ratio of infected individuals whom has ever acquired infection during the epidemic period. For this we define the local final size of epidemic  $\Psi_i^R$  within the non-commuting resident population at  $i$ -th station as

$$\Psi_i^R \equiv \frac{R_i^R(\infty)}{N_i^R} = \frac{N_i^R - S_i^R(\infty)}{N_i^R} \cong 1 - \frac{S_i^R(\infty)}{S_i^R(0)}, \quad (\text{A5})$$

and the local final size of epidemic  $\Psi_{ij}^C$  within the commuting population consisted from a commuters residing at  $i$ -th home population and working at  $j$ -th work population as

$$\Psi_{ij}^C \equiv \frac{R_{ij}^C(\infty)}{N_{ij}^C} = \frac{N_{ij}^C - S_{ij}^C(\infty)}{N_{ij}^C} \cong 1 - \frac{S_{ij}^C(\infty)}{S_{ij}^C(0)}. \quad (\text{A6})$$

Here, we have used the approximations  $S_i^R(0) \cong N_i^R$  and  $S_{ij}^C(0) \cong N_{ij}^C$ . These can be calculated in the following way. The integral forms of Eqs. (1), (4) and Eqs. (3), (6) are given as following.

$$S_i^R(t) = S_i^R(0) \exp \left\{ -\beta \int_0^t \left[ 2I_i^R(s) + \sum_k I_{ik}^C(s) \right] ds \right\} \quad (\text{A7})$$

$$S_{ij}^C(t) = S_{ij}^C(0) \exp \left\{ -\beta \int_0^t \left[ I_i^R(s) + \sum_k \left( I_{ik}^C(s) + I_{kj}^C(s) \right) \right] ds \right\} \quad (\text{A8})$$

$$R_i^R(t) = \gamma \int_0^t I_i^R(s) ds \quad (\text{A9})$$

$$R_{ij}^C(t) = \gamma \int_0^t I_{ij}^C(s) ds \quad (\text{A10})$$

Substitution of Eqs. (A7)-(A10) with  $t = \infty$  to the definition of the local final size of epidemic Eq. (A5) and Eq. (A6) yields

$$\begin{aligned} \Psi_i^R &= 1 - \exp \left\{ -\beta \int_0^\infty \left[ 2I_i^R(s) + \sum_k I_{ik}^C(s) \right] ds \right\} \\ &= 1 - \exp \left\{ -\frac{\beta}{\gamma} \left[ 2N_i^R \Psi_i^R + \sum_k N_{ik}^C \Psi_{ik}^C \right] \right\}, \end{aligned} \quad (\text{A11})$$

$$\begin{aligned} \Psi_{ij}^C &= 1 - \exp \left\{ -\beta \int_0^\infty \left[ I_i^R(s) + \sum_k \left( I_{ik}^C(s) + I_{kj}^C(s) \right) \right] ds \right\} \\ &= 1 - \exp \left\{ -\frac{\beta}{\gamma} \left[ N_i^R \Psi_i^R + \sum_k \left( N_{ik}^C \Psi_{ik}^C + N_{kj}^C \Psi_{kj}^C \right) \right] \right\}. \end{aligned} \quad (\text{A12})$$

This system of transcendental equation with  $M(M+1) \sim 100,000$  equations can be solved numerically; by recursive calculation starting from  $\Psi_i^R = 1$  and  $\Psi_{ij}^C = 1$  until all the values converges to a fixed point. Once this value  $\Psi_{ij}^C$  has been obtained, the local final size of epidemic for commuting population at each home population  $\Psi_i^H$  and each work population  $\Psi_j^W$  can be obtained in the following way.

$$\psi_i^H \equiv \frac{R_i^H(\infty)}{N_i^H} = \frac{\sum_j N_{ij}^C \psi_{ij}^C}{N_i^H} \quad (\text{A13})$$

$$\psi_j^W \equiv \frac{R_j^W(\infty)}{N_j^W} = \frac{\sum_i N_{ij}^C \psi_{ij}^C}{N_j^W} \quad (\text{A14})$$

Here the number of recovered individuals at  $i$ -th home population ( $j$ -th work population) is denoted as  $R_i^H(t) \equiv \sum_j R_{ij}^C(t)$  ( $R_j^W(t) \equiv \sum_i R_{ij}^C(t)$ ). The final size of epidemic within a total non-commuting population  $\Psi^R$  and total commuting population  $\Psi^C$  can be obtained in similar fashion from  $\psi_i^R$  and  $\psi_{ij}^C$  as

$$\Psi^R = \frac{\sum_i R_i^R(\infty)}{\sum_i N_i^R} = \frac{\sum_i N_i^R \psi_i^R}{N^R}, \quad (\text{A15})$$

$$\Psi^C = \frac{\sum_i \sum_j R_{ij}^C(\infty)}{\sum_i \sum_j N_{ij}^C} = \frac{\sum_i \sum_j N_{ij}^C \psi_{ij}^C}{N^C}. \quad (\text{A16})$$

Then the global final size of epidemic  $\Psi$  is given from  $\Psi^R$  and  $\Psi^C$  as

$$\Psi \equiv \frac{\sum_i R_i^R(\infty) + \sum_i \sum_j R_{ij}^C(\infty)}{N} = \frac{N^R \Psi^R + N^C \Psi^C}{N}. \quad (\text{A17})$$

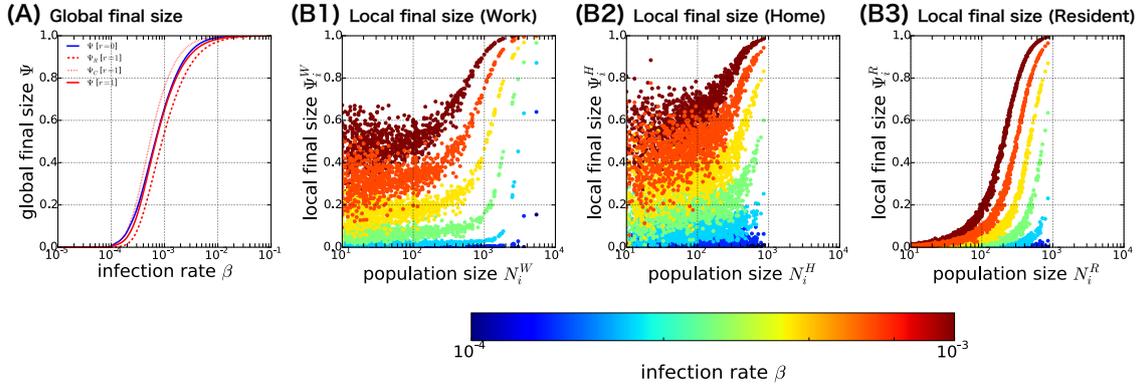
Dependence of the final size of epidemic within the non-commuting population  $\Psi^R$ , the commuting population  $\Psi^C$  and for the total population  $\Psi$  on the infection rate  $\beta$  is given in Figure S1-A. For both cases of  $r = 0$  and  $r = 1$ , when the infection rate is small, the final sizes of epidemic are negligible such that the initial extinction of disease occurs, however as the infection rate increases the global final sizes of epidemic monotonically increase until it saturates to one. The infection rates at these disease invasion thresholds agree with  $\beta_c$  obtained from the calculation of basic reproductive ratio (cf. Eq. 8). These critical infection rate  $\beta_c$  was slightly larger for  $r = 0$  ( $\beta_c = 9.210485 \times 10^{-5}$ ) compare to  $r = 1$  ( $\beta_c = 9.207523 \times 10^{-5}$ ), this can be attributed to the larger total population size for  $r = 1$ . This fact suggests that the effect of non-commuting population is minimal for the disease invasion condition, such that even though the total population is doubled the effect to the threshold value  $\beta_c$  is minimal. For  $r = 1$ , the final size of epidemic is larger for commuting population ( $\Psi^C$ ) compare to that of the non-commuting population ( $\Psi^R$ ), and the disease invasion occurs from slightly smaller infection rate. These can be ascribed to the fact that the commuting individuals have a higher risk of encountering infectious individuals at the

work population compare to the non-commuting residents. In such way, there is a quantitative difference between the results of  $r = 0$  and  $r = 1$ . However, qualitative differences are minor, therefore, throughout this study we have used  $r = 1$  in the analysis.

#### **References for supplementary materials:**

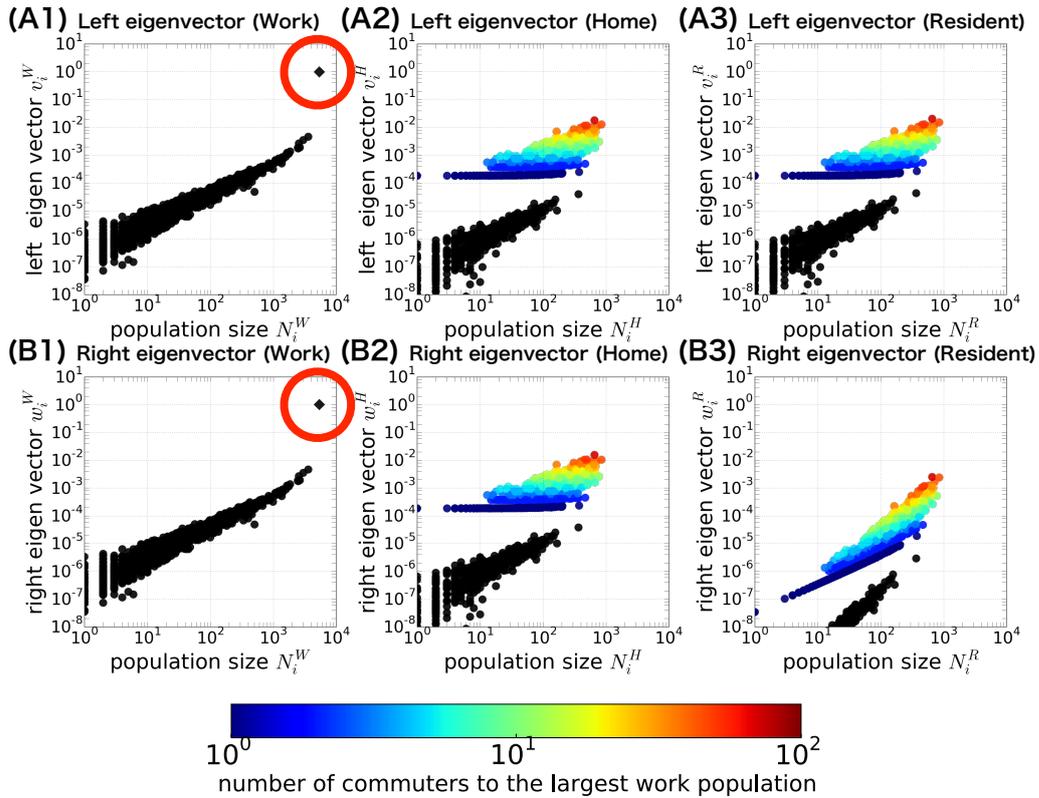
1. Diekmann O, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation* (Wiley Series in Mathematical & Computational Biology). JOHN WILEY & SON; 2000.
2. Caswell H. *Matrix Population Models*. Second. Sinauer Associates Inc.; 2001.
3. Ellner SP, Guckenheimer J. *Dynamic Models in Biology*. Princeton University Press; 2011.
4. Klemm K, Serrano MA, Eguiluz VM, Miguel MS. A measure of individual role in collective dynamics: spreading at criticality. *Sci Rep*. 2012;2: 292. doi:10.1038/srep00292

## Supporting Information S3. Supporting Figures



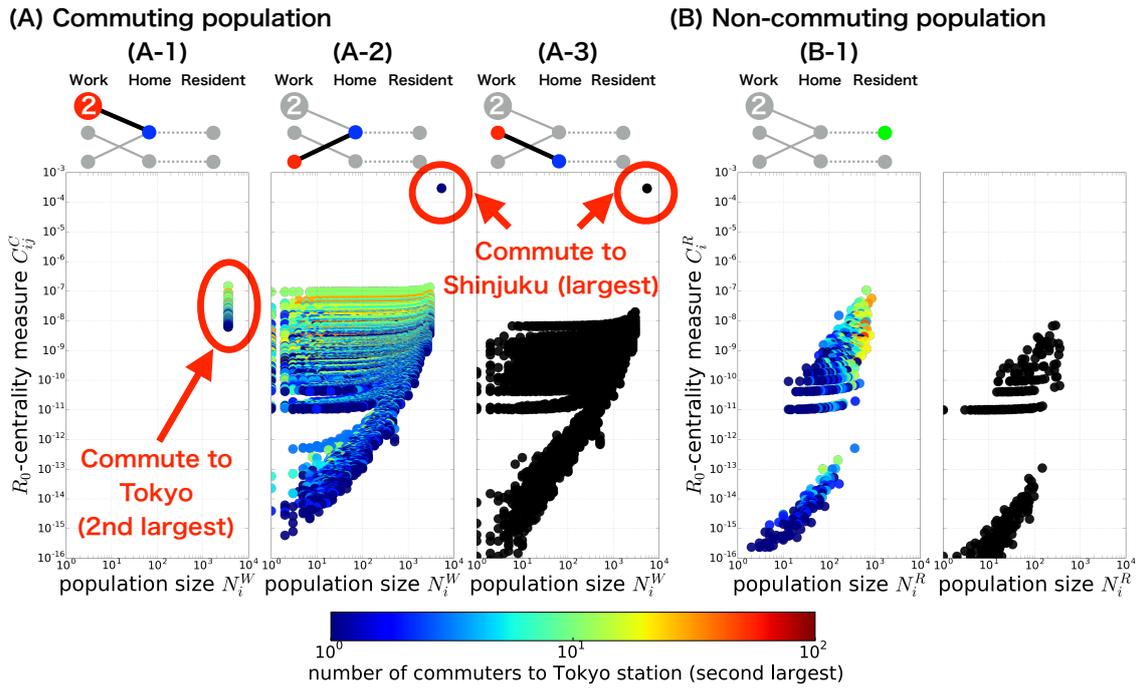
**Figure S1: The final size of epidemic for infectious disease spread in the Tokyo metropolitan area.** (A) The global final size  $\Psi$  of the epidemic for  $r = 0$  (blue line) and  $r = 1$  (red lines) are plotted against the infection rate  $\beta$  ( $r$ : ratio of non-commuting individuals to commuting individuals, see Methods section for details). For  $r = 0$ , where all the population would commute, the result for the commuting population  $\Psi^C$  (in this case the same as the result for the total population) is only present as a blue solid line. For  $r = 1$ , the type of line indicates the result for the commuting population  $\Psi^C$  (red dotted line), non-commuting population  $\Psi^R$  (red dashed line), and total population  $\Psi$  (red solid line), respectively. Note that the result for the total population with  $r = 1$  (red solid line) overlaps with the result for commuting population with  $r = 0$  (blue solid lines) and is not visible on the figure. As the infection rate  $\beta$  exceeds a threshold value, the global final size  $\Psi$  of the epidemic becomes non-zero and increases along with the infection rate, for both values of  $r$ . According to the analysis of the basic reproductive ratio, the threshold value of infection  $\beta_c$  is given as  $\beta_c = 9.210485 \times 10^{-5}$  for  $r = 0$  and  $\beta_c = 9.207523 \times 10^{-5}$  for  $r = 1$ . These values are in good agreement with the results obtained for the final size of the epidemic. (B) The local final size of the epidemic at work population ( $\Psi_i^W$ ), home population ( $\Psi_i^H$ ), and resident population ( $\Psi_i^R$ ) of each station are plotted against its local population size in figures B1, 2, and 3, respectively. The results for different infection rates are denoted by different colors, here  $r = 1$  is used for the calculation. There is a sigmoidal dependence of local final size of epidemic on its

population size, such that the local final size is small when the population size is small and as the population size becomes larger it will increase until it saturates to one at the larger limit. Here the location of the steep transition point will shift to the smaller side as the infection rate becomes larger. This point will become relevant in relation to the effect of countermeasures on the final size of epidemic (see Figure 4B, C).

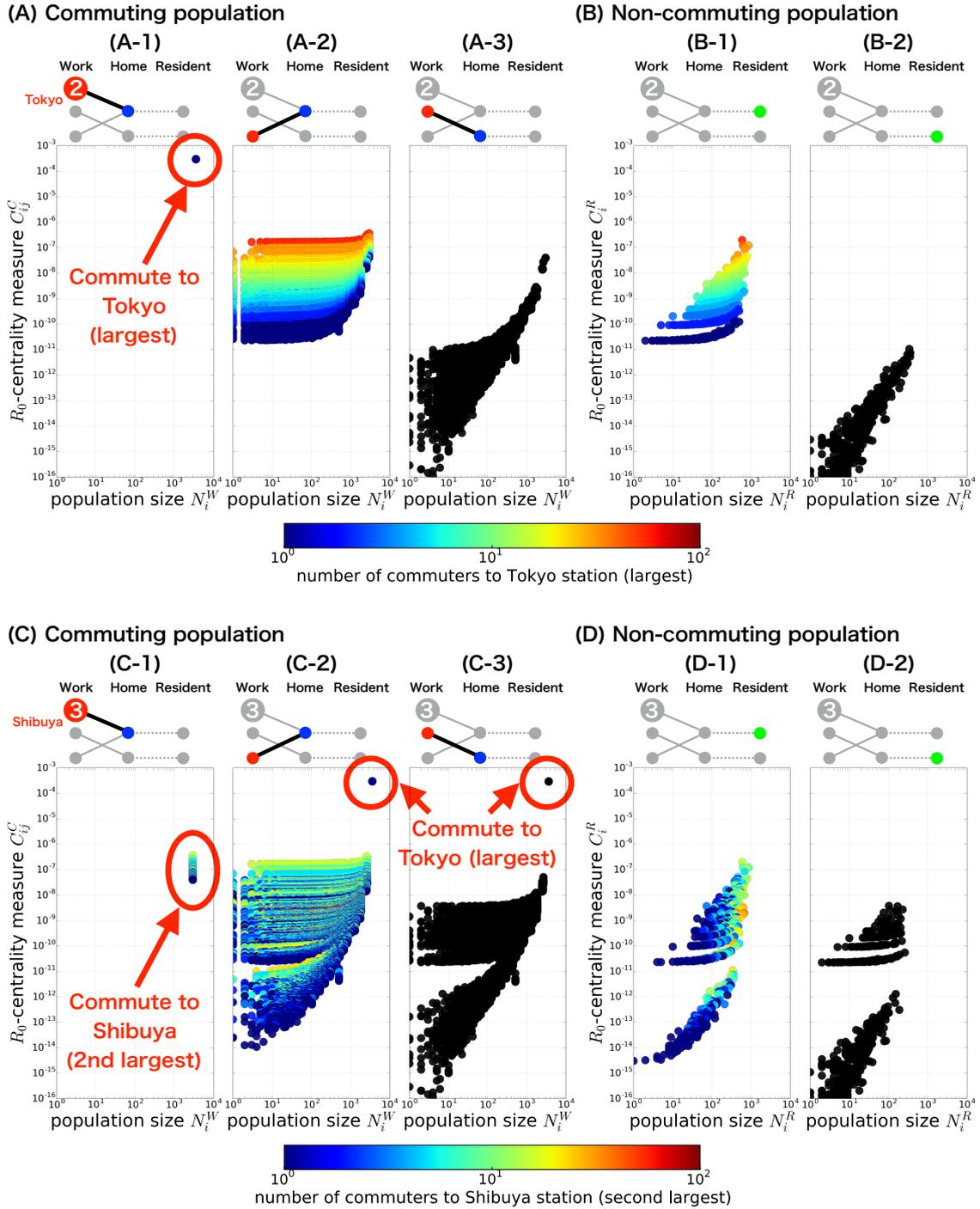


**Figure S2: The left and right eigenvectors corresponding to the dominant eigenvalue of next generation matrix for infectious disease spread in the Tokyo metropolitan area.** (A) The element of the left eigenvector ( $v_i^W$ ,  $v_i^H$ ,  $v_i^R$ ) that gives the reproductive value of infection at each local population is plotted against its local population size ( $N_i^W$ ,  $N_i^H$ ,  $N_i^R$ ), where each dot represents a single station. The “dynamic influence” introduced by Klemm et al. [4] corresponds to this value, except that they have calculated the eigenvector of the Jacobian matrix and not the next generation matrix. (B) The element of right eigenvector ( $w_i^W$ ,  $w_i^H$ ,  $w_i^R$ ) that gives the relative fraction of infected individuals at each local population in an exponentially growing phase is plotted against its local population size ( $N_i^W$ ,  $N_i^H$ ,  $N_i^R$ ), where each dot represents a single station. Results for the commuting population at each work population and home population are given in (A1, B1) and (A2, B2), respectively and the results for the non-commuting resident population at each station are given in (A3, B3). The color of each dot shows their relationship with the largest work population (Shinjuku station). Black diamonds marked with a red circle in (A1) and (B1) correspond to the largest work population, and other work populations are represented by black triangles. Each colored dot in (A2, 3) and (B2, 3) corresponds to a station that

has at least one commuter that travels to the largest work population and the color indicates the number of commuters who go there. Black dots correspond to stations with no commuters to the largest work population. For both commuting and non-commuting populations, the elements of the leading left and right eigenvectors were separated into two distinct groups, which can be interpreted from their relationship with the largest work population. The strong dependence of the  $R_0$ -centrality on the Shinjuku station originates from this characteristic.

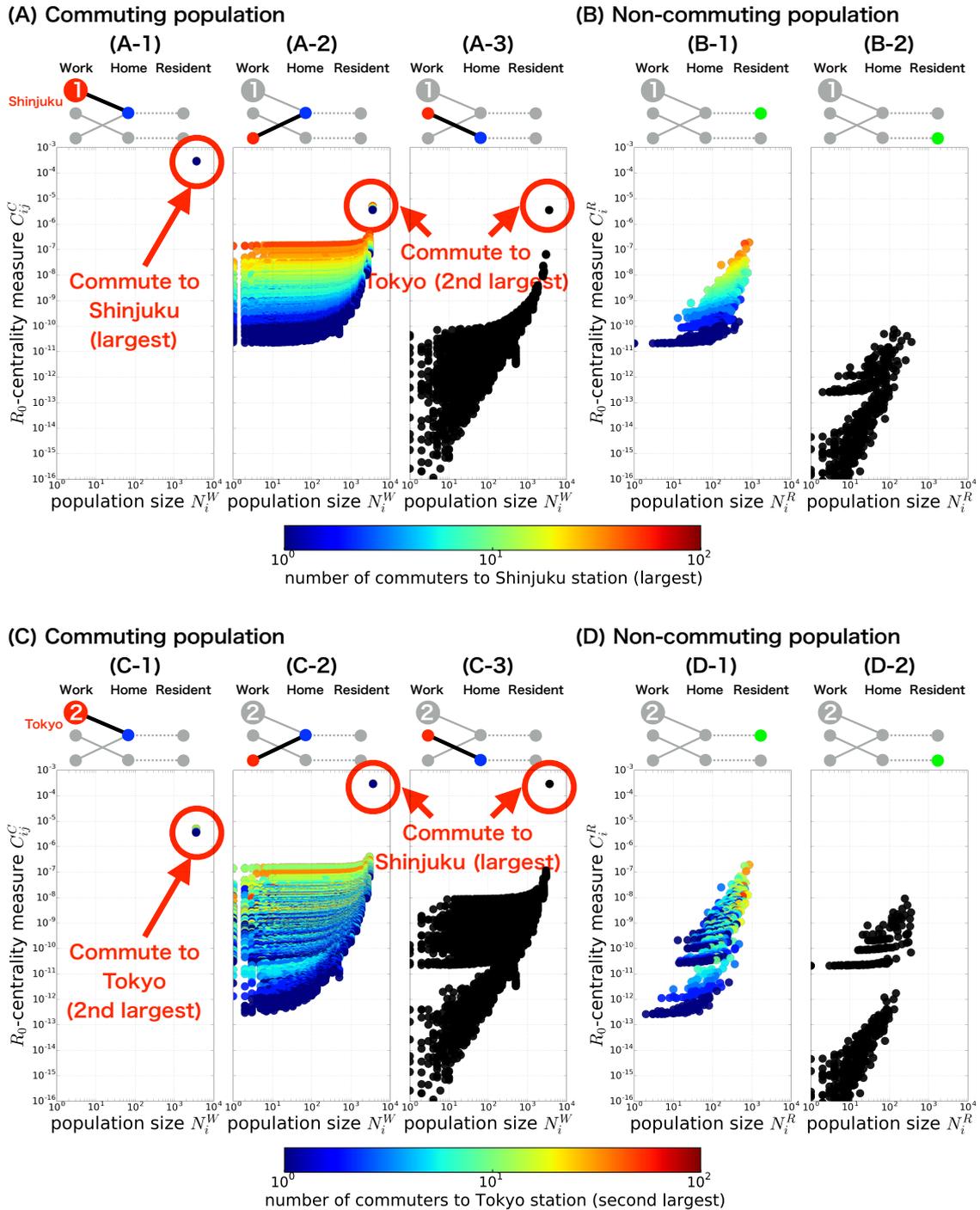


**Figure S3: The  $R_0$ -centrality for every commuting pathway and residential station in the Tokyo metropolitan area (the same data as Figure 2 presented in relation to the second largest work population). The  $R_0$ -centrality for each commuting population (each dot in Figure S3A corresponds to a single commuting pathway) and non-commuting population (each dot in Figure S3B corresponds to a single residential station) are given in accordance with the relation to the working population at Tokyo station. The schematic illustration above each panel describes its relationship. The  $R_0$ -centralities in the commuting populations (A-1), those who commute directly to Tokyo station, (A-2), those who do not commute to Tokyo station but share a common resident station with them, (A-3): neither of them, are plotted against the population size of its working population ( $N_j^W$ ). Similarly, the  $R_0$ -centralities of non-commuting population (B-1), those residing at the station area from which at least one commutes to Tokyo station, and (B-2), those residing at the station area from which no one commutes to Tokyo station, are plotted against the population size of its resident population ( $N_j^R$ ). The color of dots indicates the number of commuters to the working population at Tokyo station.**



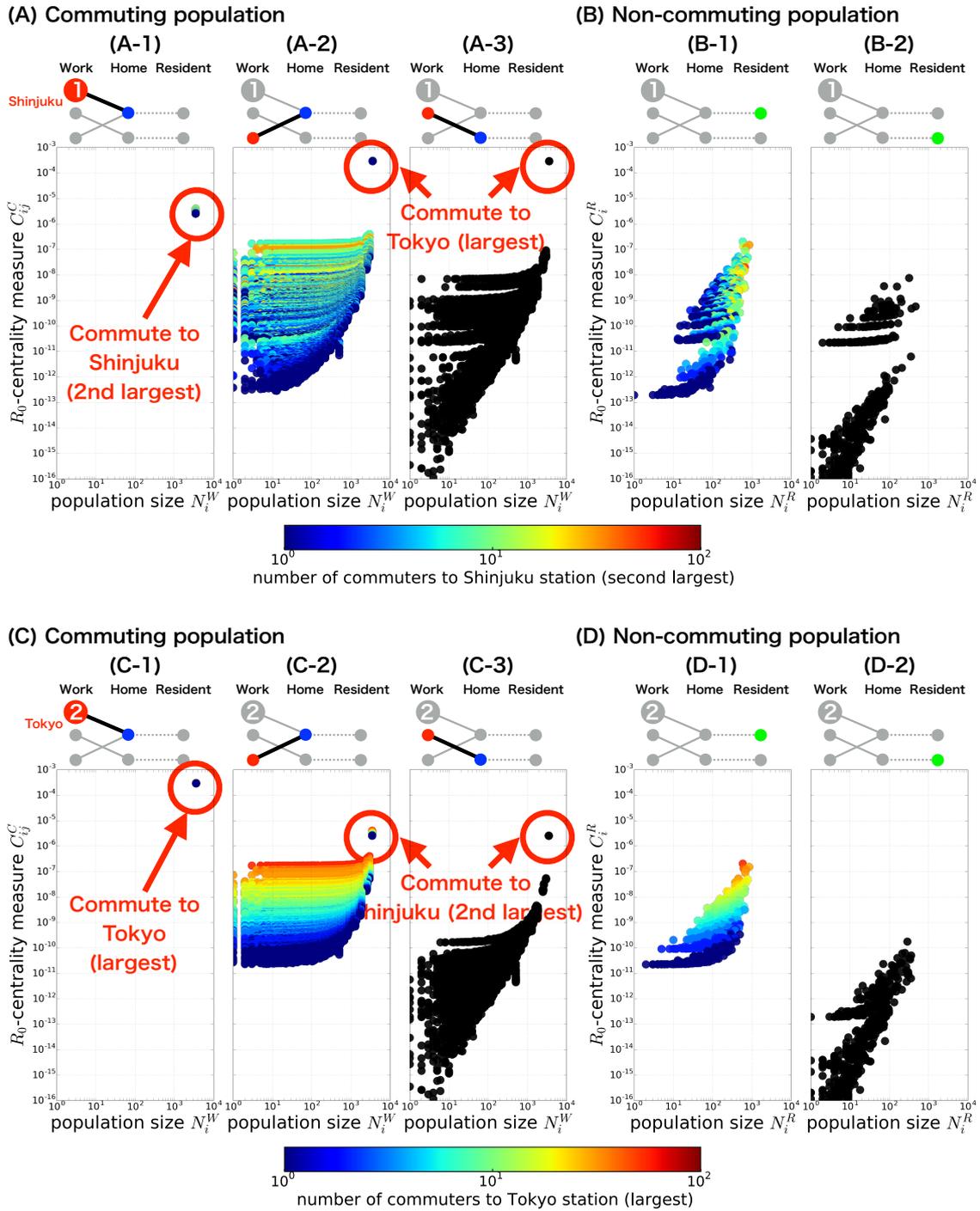
**Figure S4: The  $R_0$ -centrality for every commuting pathway and residential station in the Tokyo metropolitan area, after vaccinating/quarantining every individual from the largest working population at Shinjuku station. The  $R_0$ -centrality for each commuting population and non-commuting population after the vaccinating/quarantining all the individual from the largest working population at Shinjuku station, are given in accordance with the relation to the working population at**

Tokyo station (currently the largest susceptible work population after the removal of Shinjuku station) and Shibuya station (currently the second largest susceptible work population) are given in (A for commuting population, B for non-commuting population) and (C for commuting population, D for non-commuting population), respectively. The schematic illustration above each panel describes its relationship. The color of dots indicates the number of susceptible commuters to the working population at Tokyo station in (A, B) and to the working population at Shibuya station in (C, D).



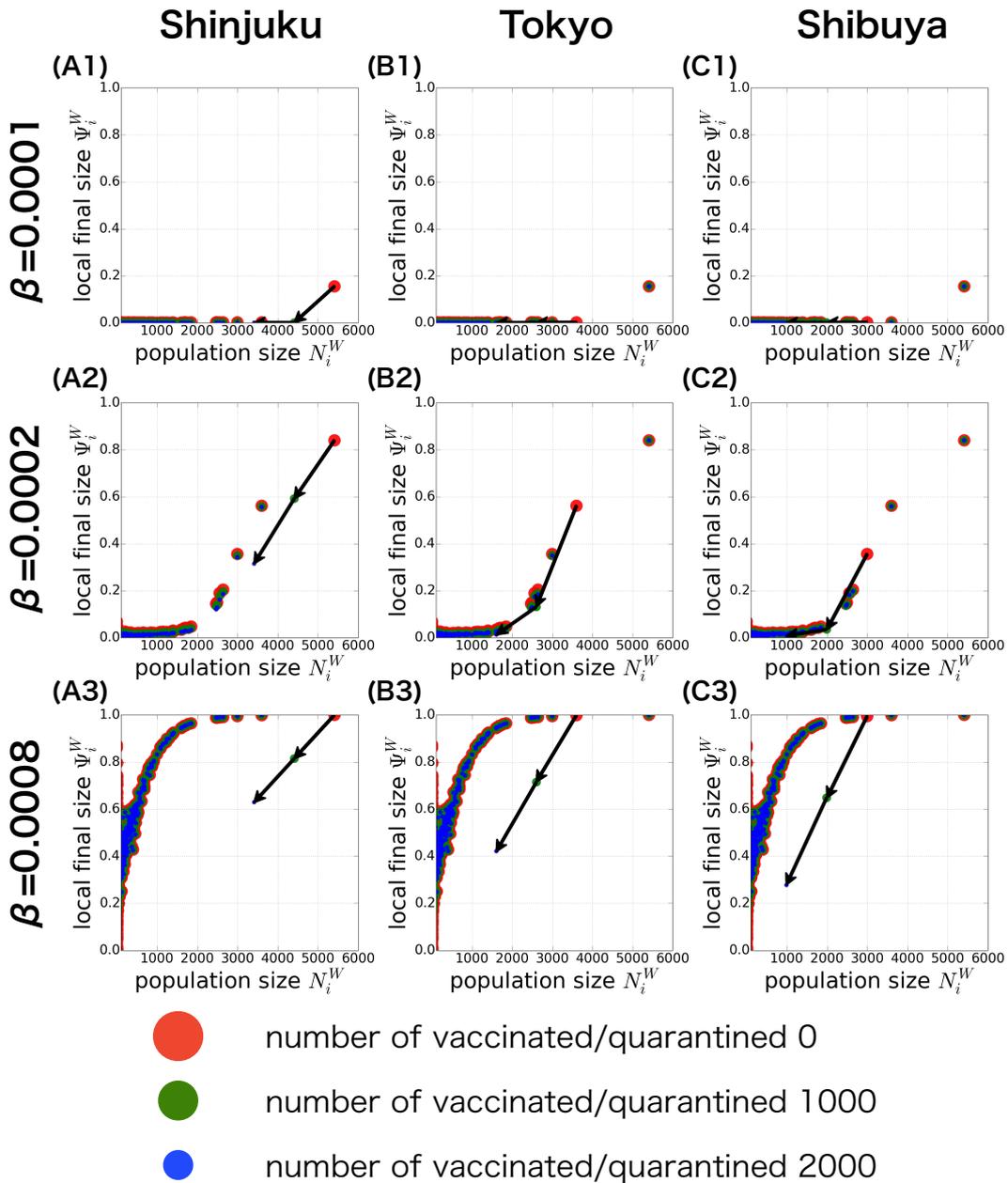
**Figure S5: The  $R_0$ -centrality for every commuting pathway and residential station in the Tokyo metropolitan area, after vaccinating/quarantining 1,700 individuals from the largest working population at Shinjuku station. The  $R_0$ -centrality for each commuting population and non-commuting population after vaccinating/quarantining 1,700 individuals from the largest working population at Shinjuku station, are given in accordance with the relation to the working population at Shinjuku station (the largest**

susceptible work population) and Tokyo station (the second largest susceptible work population) are given in (A for commuting population, B for non-commuting population) and (C for commuting population, D for non-commuting population), respectively. The schematic illustration above each panel describes its relationship. The color of dots indicates the number of susceptible commuters to the working population at Shinjuku station in (A, B) and to the working population at Tokyo station in (C, D).



**Figure S6: The  $R_0$ -centrality for every commuting pathway and residential station in the Tokyo metropolitan area, after vaccinating/quarantining 1,900 individuals from the largest working population at Shinjuku station. The  $R_0$ -centrality for each commuting population and non-commuting population after vaccinating/quarantining 1,900 individuals from the largest working population at Shinjuku station, are given in accordance with the relation to the working population at Shinjuku station (currently the**

second largest susceptible work population after vaccination) and Tokyo station (currently the largest susceptible work population after vaccination) are given in (A for commuting population, B for non-commuting population) and (C for commuting population, D for non-commuting population), respectively. The schematic illustration above each panel describes its relationship. The color of dots indicates the number of susceptible commuters to the working population at Shinjuku station in (A, B) and to the working population at Tokyo station in (C, D).



**Figure S7: The effect of countermeasures on the local final size of epidemic at each major station in the Tokyo metropolitan area.** The change in the local final size of epidemic when the vaccination/quarantine is independently applied to the working population of each major station, Shinjuku, Tokyo, and Shibuya are given in (A), (B), and (C), respectively. Here, the vaccination/quarantine is applied to the relevant population only and the other populations are kept untouched. The result for vaccinating 0 (red circle dot), 1,000 (green circle dot) and 2,000 (blue circle dot) individuals are given and the local final sizes of epidemic at each work population are plotted against

its local population size. Each panel corresponds to results for different infection rate. The sigmoidal profiles observed in Figure S1B are also evident here; for larger infection rate the transition point will shift to the smaller side. The overall shapes are not altered by the vaccination/quarantine, except for the relevant vaccinated/quarantined population. This is because the number of vaccinated/quarantined is minimal compare to the total population size (i.e., less than 1%), so the effect of vaccination/quarantine is limited to the particular population only. A black arrow denotes the decrease of local final size of epidemic at each vaccinated/quarantined population.