# Supplementary Information
## *Modeling human mobility responses to the large-scale spreading of infectious diseases*

Sandro Meloni, Nicola Perra, Alex Arenas, Sergio Gómez,
Yamir Moreno and Alessandro Vespignani

July 4, 2011

## Contents

**1 Networks analysis** **3**

**2 Global invasion threshold in metapopulation networks with origin-destination diffusion** **4**
2.1 Comparison with numerical results . . . . . 8
2.2 Size dependency analysis . . . . . 8

**3 Behavioral changes** **11**
3.1 Departure probability . . . . . 11
3.2 Rerouting . . . . . 12
3.3 Subpopulations affected by an outbreak as a function of $R_0$ when the rerouting mechanism is active . . . . . 13
3.4 Results . . . . . 15

**4 Model's Implementation** **16**
4.1 Internal nodes' dynamics . . . . . 16
4.2 Numerical simulations . . . . . 18

**5 Sensitivity Analysis** **18**
5.1 Effects of recovery rate $\mu$ . . . . . 19
5.2 Effects of the parameter $\tau$ accounting for the time spent at destination . . . . . 19
5.3 Effects of the mobility rate of infected individuals $\lambda^I$ . . . . . 19
5.4 Effects of rerouting intensity $h$ . . . . . 19
5.5 Effects of latency rate $\gamma$ in SEIR model . . . . . 19

Metapolulation models are a relevant framework to model spreading of infectious diseases between fragmented but coupled subpopulations. Each of these subpopulations is represented as a node in a network with a certain number of connections with others. Individuals diffuse following the paths defined by the links that connect subpopulations. Inside each subpopulation, individuals are assumed to be well-mixed, so that the evolution of the epidemic is determined by the reproductive number $R_0$ of the disease. In metapopulation systems this quantity is not sufficient to understand if a macroscopic fraction of subpopulations will experience an outbreak. The diffusion and coupling between each node must be considered. Using these arguments a new quantity is introduced: the global invasion threshold. This is in general a function of the basic reproductive number, the connectivity patterns, diffusion rates and mobility patterns. Recently, analytical solutions have been found for Markovian diffusion processes [1]. In this paper we consider a more realistic scenario for individuals' mobility. People have a home, and when they travel, they also go to a specific, not random destination [2].

In our study, we have first analytically evaluated the global invasion threshold for synthetic networks considering explicitly origin and destination matrices. We next added even more realism considering real networks, namely, the worldwide air transportation network, and different mechanisms that account for self-initiated behavioral changes that are induced by the spreading of the epidemic among different subpopulations. In doing so, we have focused on spontaneous behavioral changes due the evolution and awareness of the epidemic. For instance, when an epidemic outbreak takes place, people might minimize their contacts as a way to reduce the risk to get sick. Besides, it has also been documented that one of the most common risk aversion strategies consists of staying at home, i.e., not to engage in a travel if there is the perception (or certainty) that at destination the number of infected individuals is high. Another behavioral reaction is encountered when travels are not cancelled, but individuals decide instead to change their itinerary avoiding places in which the epidemic incidence is perceived as high. These two possible risk aversion scenarios are considered in our study through the implementation of two mechanisms accounting for such behavioral changes in a metapopulation framework. In this supplementary information, we give more details of the results presented in the main text. In particular, we first describe in details the synthetic and real world networks used in the analysis. Then we present the analytical derivation of the global invasion threshold with a simple non-markovian diffusion scheme with mobility patterns given by origins and destinations. Finally, we move to the numerical details and sensitivity analysis of the model in which real networks and behavioral changes are introduced.

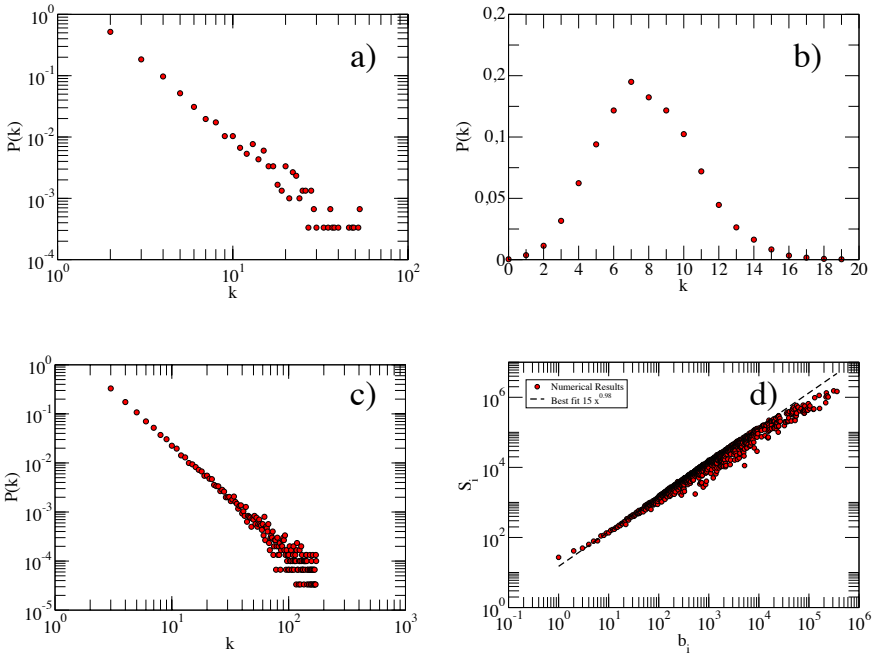


Figure 1: Degree distributions of the generated synthetic networks and correlations between nodes strength and betweenness centrality in the ATN. (a) Degree distribution $P(k)$ of the uncorrelated scale free network made up of $N = 3 \cdot 10^3$ subpopulations used as a substrate in our analysis. (b) Degree distribution $P(k)$ of a random network generated using the Erdös and Rényi model. The network is of the same size and average degree of that depicted in panel (a). (c) Degree distribution $P(k)$ of the uncorrelated scale free network with $N = 3 \cdot 10^4$ subpopulations used in the size dependency analysis. (d) Correlation between the observed traffic strength $s_i$ of each node $i$ in the ATN and the betweenness centrality values $b_i$ produced by a the shortest path routing algorithm assumed in the simulations.

# 1 Networks analysis

Before starting our analysis let us present the networks we used as substrate to simulate real world populations. First, we focus on synthetic uncorrelated scale free networks that offer a controlled test-bench for the analytical treatment of the model and then we move to a realistic scenario considering the so-called Air Transportation Network (ATN). Synthetic scale free networks were generated following the recipe in [3]. In order to resemble as much as possible the ATN, we use a network of similar size $N = 3 \cdot 10^3$. One might argue that

this is a too short network size, as it is well-known that the Mean-Field approximation in quenched networks is valid for very large systems sizes. For finite networks, finite-size corrections might be needed.

Our results show that for a synthetic system of a size equivalent to that of the ATN, the analytical results nicely fit numerical simulations. This is because for $N = 3 \cdot 10^3$, the differences between homogeneous and heterogeneous networks already manifest (see Fig. 1a and Fig.1b). However, in order to test the effects of the network size in the results shown below, we have also built a larger scale free network made up of $N = 3 \cdot 10^4$ subpopulations. Figure 1c depicts the corresponding degree distribution.

Another assumption of our model that is related to the connection between synthetic and real networks has to do with the way individuals move through the networks. Specifically, we assume that individuals move from their home to the destination following the shortest path connecting them. In Fig.1d, we compare the observed nodes' strengths $s_i$ from the ATN (i.e., real fluxes between any two given nodes) and the betweenness centrality $b_i$ of that $i$ when a shortest path routing algorithm is implemented. The linear association between the two sets of data doubtlessly indicate that the latter is indeed a quite good proxy for individual mobility.

# 2 Global invasion threshold in metapopulation networks with origin-destination diffusion

Let us consider a network $\mathbf{G}(V, M)$ with $V$ nodes, $M$ links and degree distribution $P(k)$. Each node $i$ of the network is considered a subpopulation with $N_i$ individuals. We set the subpopulation size proportional to its degree. Next, assume a diffusion process in which each individual in a node $i$ (origin) travel to another node $j$ (destination) of the network with a probability $\lambda$. The destination choice $j$ is selected with a probability proportional to the subpopulation size $N_j$. Once individuals engage in a travel, the shortest path is selected among all the possible paths connecting the origin and the destination of the travel. Besides, we also consider that individuals come back to their home subpopulations after they reach their destinations. A standard convenient representation of the system is provided by quantities defined in terms of the degree $k$:

$$N_k = \frac{1}{V_k} \sum_{i|k_i=k} N_i. \tag{1}$$

Let us consider that an individual of a subpopulation of degree $k$ gets some infectious disease characterized by a reproductive number $R_0 > 1$. Let us define $D_k^0$ as the number of diseased subpopulations of degree $k$ at generation $0$. In the early stage, the number of diseased subpopulations is small, thus, we can study the evolution of this number using a tree-like approximation (where no correlations are considered) relating $D_k^n$ with $D_k^{n-1}$. The average number of infected individuals in the class of degree $k$ during the evolution of the epidemic

is $\alpha N_k$. The parameter $\alpha$ depends on the specific disease. Each infected individual stays in the infectious state for an average time $\mu^{-1}$. Then the number of infected people circulating through the network after $n-1$ generations is:

$$\omega^{n-1} = \frac{\lambda\alpha}{\mu} \sum_{k'} D_{k'}^{n-1} N_{k'} \tag{2}$$

The number of infected individuals that will pass through a subpopulation of degree $k$ will be a fraction of Eq. (2) proportional to the topological betweenness (in general, it is proportional to the algorithmic betweenness, but given that individuals are following the shortest path, it coincides with the topological one in our case). This measure is defined as:

$$b(i) = \sum_{\substack{j,l=1,n \\ i\neq j\neq l}} \frac{\mathcal{D}_{jl}(i)}{\mathcal{D}_{jl}}, \tag{3}$$

where $\mathcal{D}_{jl}$ is the total number of shortest paths from $j$ to $l$ and $\mathcal{D}_{jl}(i)$ is the number of shortest paths from $j$ to $l$ that goes through $i$. The latter quantity also measures the centrality of a node assuming a diffusion scenario in which travelers go through the shortest paths. We can then write:

$$\gamma_k^{n-1} = \frac{b_k}{b_{tot}} \omega^{n-1}, \tag{4}$$

where $b_{tot}$ is the sum of all the betweenness of the nodes. For the $n^{th}$ generation we have:

$$D_k^n = V_k \left(1 - \frac{D_k^{n-1}}{V_k}\right) \left[1 - R_0^{-\gamma_k^{n-1}}\right], \tag{5}$$

where the second factor on the right is the probability that the subpopulation is not already seeded by infected individuals and the last is the probability that the new seeded subpopulation will experience an outbreak. In the early time and for $R_0 \sim 1$ we can approximate the last expression considering:

$$\frac{D_k^{n-1}}{V_k} << 1, \tag{6}$$

and

$$1 - R_0^{-\gamma_k^{n-1}} \sim (R_0 - 1)\gamma_k^{n-1}, \tag{7}$$

obtaining:

$$D_k^n = (R_0 - 1) V_k \gamma_k^{n-1} = (R_0 - 1) \frac{\lambda\alpha}{\mu} V_k \frac{b_k}{b_{tot}} \sum_{k'} D_{k'}^{n-1} N_{k'}. \tag{8}$$

Considering at the equilibrium:

$$N_k = \frac{k}{\langle k \rangle} \bar{N}, \tag{9}$$

where $\bar{N} = \sum_k P(k)N_k$ is the average subpopulation size, we get:

$$D_k^n = (R_0 - 1)\frac{\lambda\alpha}{\mu}\bar{N}V_k\frac{b_k}{b_{tot}}\frac{1}{\langle k\rangle}\sum_{k'} D_{k'}^{n-1}k'. \tag{10}$$

Let us define now $\Theta^n = \sum_k D_k^n k$, then we have:

$$\Theta^n = (R_0 - 1)\frac{\lambda\alpha}{\mu}\bar{N}\frac{\Theta^{n-1}}{\langle k\rangle}\sum_k V_k k\frac{b_k}{b_{tot}}. \tag{11}$$

The last term needs can be further developed as:

$$\sum_k V_k k\frac{b_k}{b_{tot}} = \frac{V\sum_k P(k)kb_k}{V\sum_{k'} P(k')b_{k'}}. \tag{12}$$

Considering now $b_k \sim k^\eta$ one is left with:

$$\Theta^n = (R_0 - 1)\frac{\lambda\alpha}{\mu}\bar{N}\frac{1}{\langle k\rangle}\frac{\langle k^{1+\eta}\rangle}{\langle k^\eta\rangle}\Theta^{n-1}. \tag{13}$$

We finally get the global invasion threshold as:

$$R^* = (R_0 - 1)\frac{\lambda\alpha}{\mu}\bar{N}\frac{1}{\langle k\rangle}\frac{\langle k^{1+\eta}\rangle}{\langle k^\eta\rangle}. \tag{14}$$

We can write the threshold condition for the mobility rate:

$$\lambda\bar{N} \geq \frac{\langle k^\eta\rangle}{\langle k^{1+\eta}\rangle}\frac{\langle k\rangle\mu}{\alpha}(R_0 - 1)^{-1}. \tag{15}$$

These last two expressions are the crucial quantities, and give the conditions for a global outbreak. It is important to remind that in metapopulation networks the condition $R_0 > 1$ for each subpopulation is not enough to infer whether a finite number of subpopulations will be affected by the disease. The diffusion process must be considered, and it defines the form and value of the invasion threshold. The previous arguments are valid for the case in which:

$$\mu^{-1} \gg \bar{l}v^{-1}, \tag{16}$$

where $\bar{l}$ is the average distance between nodes and $v^{-1}$ is the traveling *speed* of individuals that we set as $1$ node per time step. In other words, this means that sick individuals are infectious for a large enough time compared with the time it takes for individuals to complete their travel. This is a necessary condition for Eq. (4) to be valid. We can easily get an expression in the other limit:

$$\mu^{-1} \ll \bar{l}v^{-1}, \tag{17}$$

in this case each individual will be infectious for a time window smaller enough as to infect only the nearest neighbors. In this case the spreading can be thought

of as a Markovian process and all the analytics is the same reported in Ref. [1]. The expression for the number of infected nodes at the $n^{th}$ generation will be then:

$$D_k^n = \sum_{k'} D_{k'}^{n-1}(k'-1)\left[1 - R_0^{-\lambda_{k'k}}\right] P(k|k')\left(1 - \frac{D_k^{n-1}}{V_k}\right), \tag{18}$$

using the usual approximations (6), (7) and assuming that degree correlations can be neglected we have:

$$D_k^n = \frac{(R_0-1)}{\langle k\rangle} kP(k) \sum_{k'} D_{k'}^{n-1}(k'-1)\lambda_{k'k}. \tag{19}$$

Considering an homogeneous diffusion we can write:

$$\lambda_{k'k} = \frac{\lambda}{k'}\frac{\alpha N_{k'}}{\mu}, \tag{20}$$

where $N_{k'} = \frac{k'}{\langle k\rangle}\bar{N}$. Using this expression we get:

$$D_k^n = \frac{\lambda\alpha}{\mu}\frac{(R_0-1)\bar{N}}{\langle k\rangle^2} kP(k) \sum_{k'} D_{k'}^{n-1}(k'-1). \tag{21}$$

Let us defined $\Theta^n = \sum_k D_k^n(k-1)$, after multiplying both sides for $(k-1)$ and after summing over all $k$ we get:

$$\Theta^n = \frac{\lambda\alpha}{\mu}\frac{(R_0-1)\bar{N}}{\langle k\rangle^2}\Theta^{n-1} \sum_k k(k-1)P(k), \tag{22}$$

or

$$\Theta^n = \frac{\lambda\alpha}{\mu}\frac{(R_0-1)\bar{N}}{\langle k\rangle^2}\Theta^{n-1}(\langle k^2\rangle - \langle k\rangle). \tag{23}$$

We finally got an expression for the global invasion threshold:

$$R^* = \frac{\lambda\alpha}{\mu}(R_0-1)\bar{N}\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle^2}, \tag{24}$$

and for the mobility rate

$$\lambda\bar{N} \geq \frac{\mu}{\alpha(R_0-1)}\frac{\langle k\rangle^2}{\langle k^2\rangle - \langle k\rangle}. \tag{25}$$

## 2.1 Comparison with numerical results

To compare the analytical approach with the numerical results we choose as substrate an uncorrelated scale free network generated according to the uncorrelated configuration model with $\gamma = 2.5$ and $N = 3000$. First of all we tested the assumption made in Eq. (9), in which the number of individuals $N_k$ at nodes of degree $k$, at the equilibrium, is proportional to $k$. To do so, we start the simulation with a population of $N_i \simeq 1000$ in each node, wait until the traffic equilibrium has been reached, and finally we collect the values of $N_i$. Fig. (2)a shows the values of $N_k$ as function of degree $k$, justifying our assumption. In order to calculate the critical mobility rate $\lambda_c$ we use Eq. (15) and $\lambda_c$ reads as:

$$\lambda_c = \frac{1}{\bar{N}} \frac{\langle k^\eta \rangle}{\langle k^{1+\eta} \rangle} \frac{\langle k \rangle \mu}{\alpha} (R_0 - 1)^{-1}. \tag{26}$$

Thus, we need to know the specific value of $\eta$ in the chosen network and fix a value for $R_0$. To obtain an estimate for $\eta$ we compute the value of the betweenness $b_i$ for each node $i$ and coarse grain it by degree classes $k$ as,

$$b_k = \frac{1}{V_k} \sum_{i|k_i=k} b_i. \tag{27}$$

Note that to evaluate $b_i$ of each node we decide to make a run of the simulation and register the number of packets that pass through a link over a very long period of time. In this way, the values of $b_i$ are more precise and closer to the actual dynamics. Fig. (2)b shows $b_k$ as function of the degree classes and the fit for $\eta$ gives a value of $\eta = 1.51$.

Now we have all the ingredients to calculate the critical value $\lambda_c$ and compare it with the numerical results. First we calculate the mean degree of the network $\langle k \rangle = 4.0$ and then the $\eta^{th}$ moment of the degree distribution obtaining $\langle k^\eta \rangle = 10.93$ and finally the $(1+\eta)^{th}$ moment $\langle k^{1+\eta} \rangle = 158.60$. Considering $R_0 = 1.5$ and that $\bar{N} = 1000$, and substituting $\alpha$ in Eq. (26) one gets

$$\lambda_c = \frac{1}{\bar{N}} \frac{\langle k^\eta \rangle}{\langle k^{1+\eta} \rangle} \frac{\langle k \rangle \mu R_0^2}{2(R_0 - 1)^2} = \frac{1}{1000} \frac{10.93}{158} 4.0 \cdot 0.04 \cdot \frac{1.5^2}{2 \cdot 0.5^2} = 0.0000496. \tag{28}$$

In Fig. 2c we show the good agreement obtained when comparing the numerical simulations of the model and the analytical prediction for the global invasion threshold. Besides, a similar calculation but for $\mu = 0.5$ gives the second limit Eq. (17), which is compared in Fig. 2d with numerical simulations, showing again a good agreement.

## 2.2 Size dependency analysis

As previously argued, one would expect finite size corrections to the Mean-Field approach for small systems. In order to address this issue, we consider a larger synthetic scale free network made up of $N = 3 \cdot 10^4$ nodes. Also for this

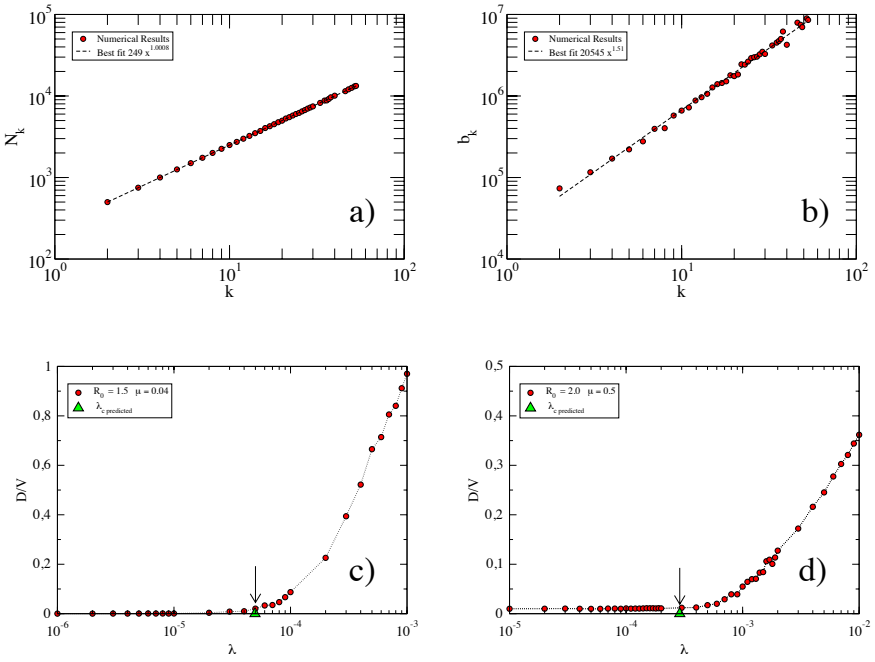


Figure 2: Panel (a) represents $N_k$ as a function of the degree classes $k$ and the best fit, which leads to a linear scaling. In panel (b), we show $b_k$ as a function of the degree classes $k$ and the best estimated value for $\eta$ is $1.51$. In panels (c) and (d), we show the fraction of infected subpopulations in the null model as a function of the mobility rate $\lambda$ for two different values of $\mu$ such that $\mu^{-1} \gg \bar{l}v^{-1}$ (c) and $\mu^{-1} \ll \bar{l}$ (d).

case we have checked the assumption made in Eq. 9 and estimated the value of $\eta$ from the relation between $k$ and $b_k$. In Fig. 3a we show the scaling between the population on nodes of degree $k$ and $k$. Again, the scaling turns out to be linear. Moreover, Fig.3b represents the dependency of $b_k$ with $k$, from which a value of $\eta = 1.161$ is obtained.

We have also tested the robustness of our derivation when infected individuals move at a lower rate than other individuals, which is a realistic ingredient. In this case we simply need a rescaling of Eq. 26. If, for simplicity, we assume $\lambda^I = \lambda/2$ we obtain:

$$\lambda_c = \frac{2}{\bar{N}} \frac{\langle k^\eta \rangle}{\langle k^{1+\eta} \rangle} \frac{\langle k \rangle \mu}{\alpha} (R_0 - 1)^{-1}. \tag{29}$$

Now considering the mean sub-population size $\bar{N} = 1000$ and $R_0 = 1.5$ we can

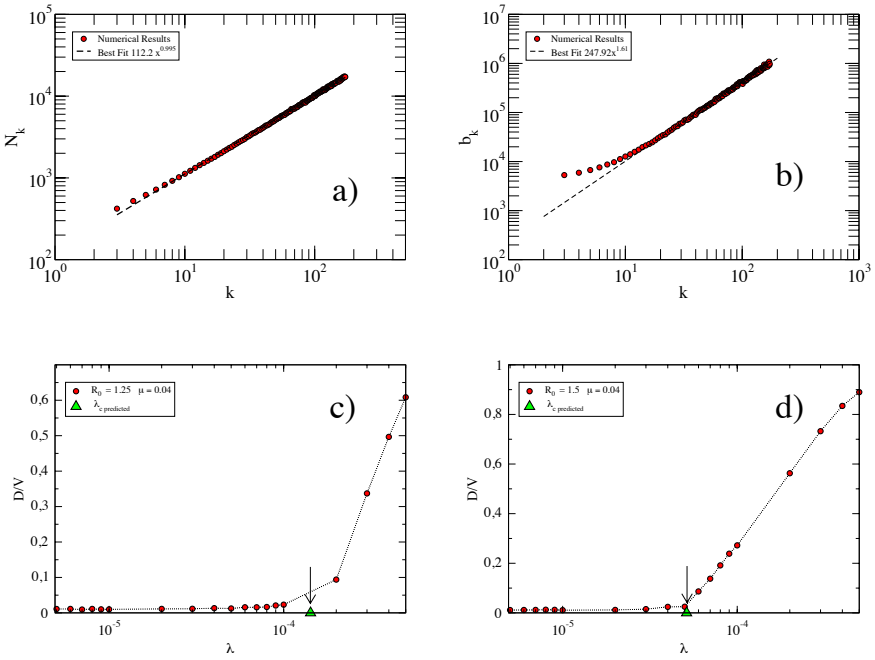


Figure 3: Panel (a) represents $N_k$ as a function of the degree classes $k$. The best fit leads to a linear scaling for the $N = 3 \cdot 10^4$ uncorrelated scale free network. In panel (b), we show $b_k$ as a function of the degree classes $k$, from which a best estimate gives $\eta = 1.51$. Panels (c) and (d) depict the fraction of infected subpopulations in the null model as a function of the mobility rate $\lambda$ for two different values of $R_0$, $R_0 = 1.25$ (c) and $R_0 = 1.5$ (d).

calculate the invasion threshold value as:

$$\lambda_c = \frac{2}{\bar{N}} \frac{\langle k^{\eta} \rangle}{\langle k^{1+\eta} \rangle} \frac{\langle k \rangle \mu R_0^2}{2(R_0 - 1)^2} = \frac{2}{1000} \frac{61.91}{3817.69} 8.78 \cdot 0.04 \cdot \frac{1.5^2}{2 \cdot 0.5^2} = 0.0000513068. \tag{30}$$

Figures. 3c and 3d show that also in this case we get a good agreement between numerical simulations and the analytical prediction.

# 3 Behavioral changes

During the evolution of an infectious disease, one usually finds that due to risk perception, people change their traveling habits. Individuals can decide to stay at home, to avoid crowded places, to reschedule trips, holidays etc.. All realistic epidemic models are based on structured data of human activities [5, 6, 7]. During catastrophic events the system can be driven out of equilibrium invalidating the predictive power of existing models. Therefore, behavioral changes should be in principle considered in any realistic attempt to model the dynamics of global infectious diseases. In the recent literature, a few attempts to model behavioral changes have been done [8, 9, 10, 11, 12, 13, 14] but any of these tackle the effect of social disruption inside a metapopulation framework. Motivated by this, we introduce different mechanisms that might model spontaneous self-initiated behavioral responses to the presence of a disease. In our full model with such changes incorporated, origin-destination matrices are still considered. Furthermore, with the aim of being more realistic, we use in our simulations the worldwide air transportation network in which each node/subpopulation $i$ represents an airport and for each connected pair $(i, j)$ a weight $\omega_{ij}$ (number of passengers in that route ) is assigned. The number of individuals in each subpopulation is set proportional to the strength:

$$N_i = \sum_j \omega_{i,j}. \tag{31}$$

As we have discussed in the previous sections, each individual will chose the destination among all possible subpopulations with a probability that is proportional to the subpopulations sizes. In this way the known heterogeneity of the traffic is explicitly considered. Let us now imagine that an individual from node $i$ selects as destination subpopulation $j$ and let $d_{ij}$ be the traveled distance in terms of hops needed to reach the chosen destination. In the previous model we assumed that once the destination is reached, each individual goes back immediately. Instead in this second more realistic implementation we assume that each traveler stay $\tau$ time steps at destination. We extract the waiting times $\tau$ from a uniform distribution with mean $d_{ij}$, so that the longer the trip, the larger the stay. Moreover, we also assume that the probability that an infected individual will travel is half of the same quantity for healthy people, as infected people once aware of the illness usually refrain from traveling. These elements constitute all together the new baseline of our model. In the next sections we describe the different mechanisms we use to model behavioral changes.

## 3.1 Departure probability

During the H1N1 pandemic in 2009, especially in the early stage, a big drop in the number of travelers to (and within) Mexico was registered [15]. A first plausible mechanism to model behavioral changes is obtained changing the probability of departure according to the stage of the disease at a given destination. We can thus assume that individuals might decide to postpone their

trips. Mathematically this behavior can be modeled as:

$$\lambda \to \lambda_{ij} = \lambda \left[ 1 - \frac{I_j(t)}{N_j(t)} \right]. \tag{32}$$

The expression above considers that the probability that each individual will travel is not anymore constant but a function of the epidemic incidence at destination. At the beginning of the spreading process, when the number of infected individuals is small, the mobility is given by $\lambda$ for all possible destinations $j$, since the second factor above is close or equal to $1$. However, during the evolution of the disease, as soon as the number of infected individuals increases, the departure probability starts to change from place to place and the mechanism becomes effective.

## 3.2 Rerouting

Let us assume that a traveler from subpopulation $i$ has as destination subpopulation $j$, and that a node $m$ is in the shortest path between its origin and destination. Let us also suppose that subpopulation $m$ is experiencing a severe outbreak. The individual could decide to travel anyway but changes the route going through another, maybe longer but less risky alternative path. We modeled this kind of behavior by introducing a cost function:

$$c_m(t) = h\delta_m + (1-h)\frac{I_m(t)}{N_m(t)}, \tag{33}$$

where the parameter $h$ is defined in the closed interval $[0, 1]$ and $\delta_k$ can assume three values $[-1, 0, 1]$. $-1$ is associated to the shortest path (the individual will be one hop closer to its destination), $0$ to a new path which does not change the current distance to the destination, and $1$ otherwise (the individual will be one hop farther to its destination). At each time step, each individual that is traveling decides the next node to move to by minimizing the cost function Eq. (33), unless the next move leads to its destination. Other functional forms for Eq. (33) could also be defined, however, Eq. (33) is a simple proxy for an stochastic choice based on the minimum information available.

It is worth noticing that the parameter $h$, although defined in the interval $[0, 1]$ can take on only a small subset of meaningful values. If $h$ is too small, the traveler essentially moves through the network following the landscape of epidemic incidence, as no information of its destination is taken into account when deciding where to move. This is a highly unrealistic situation that therefore sets a lower bound ($> 0$) to $h$. Similarly, one can easily show that $h$ is also bounded from above. Although $h = 1$ mathematically corresponds to the limit of shortest path, this limit is obtained well before. Admittedly, one can show that in order for a traveler to go through a path one hop farther than the destination the following condition must be satisfied

$$\frac{I_-(t)}{N_-(t)} - \frac{I_+(t)}{N_+(t)} > \frac{2h}{1-h}, \tag{34}$$

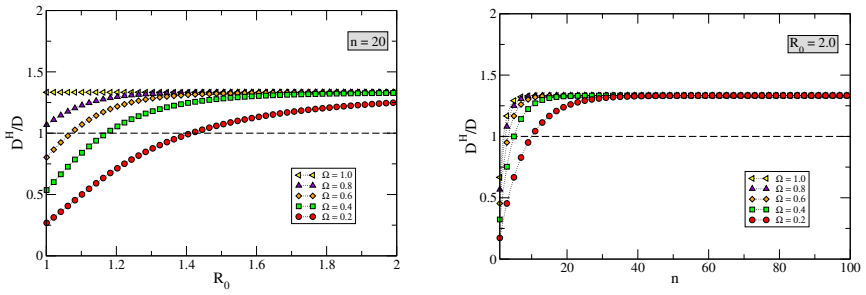


Figure 4: Numerical solution of the ratio between the number of diseased subpopulation with and without rerouting, Eq. (39), for different values of $n$, $R_0$ and $\Omega$ as indicated in the panels. The other parameters have been set to $l = 3$ and $\Delta l = 1$. The lines are a guide to the eye and mark the values of $R_0$ and $n$ at which the crossover takes place.

where $\frac{I_-}{N_-}$ and $\frac{I_+}{N_+}$ are the densities of infected individuals at subpopulations one hop closer and one hop farther from the traveler's destination, respectively. The same argument leads to the following condition with respect to the possibility of going through a path that does not change the current distance to destination:

$$\frac{I_-(t)}{N_-(t)} - \frac{I_=(t)}{N_=(t)} > \frac{h}{1-h}, \tag{35}$$

where $\frac{I_=}{N_=}$ is the density of infected individuals at a subpopulation which is at the same distance of the traveler's destination. We have checked that our results for the model including behavioral changes are qualitatively the same for different values of $h$ in the interval $[0.05, 0.2]$ (beyond $h = 1/3$, no differences with respect to the shortest path results are obtained, which means that this limit has been reached at $h = 1/3$ as expected from Eq. (34)).

### 3.3 Subpopulations affected by an outbreak as a function of $R_0$ when the rerouting mechanism is active

When the $h$ mechanism is active, the number of individuals going through a node in the shortest path between an origin and a destination decreases due to the rerouting of individuals induced by the risk aversion mechanism. This implies that a smaller number of individuals goes through that node. This mechanism defines an interesting phenomenology whose effect on the number of diseased subpopulations is non-linear and depending on the reproductive number $R_0$ and the number of traveling individuals $n$.

Let us consider that $n(\lambda)$ traveling individuals starts at subpopulation $i$ with destination in the subpopulation $j$. Besides, consider that one node $r$ in

the path connecting both subpopulations is infected. In the case without any behavioral responses the shortest path is selected and let us consider its length as $l$. The number of new diseased subpopulations on the way will be:

$$D^{new} = l(1 - R_0^{-n}), \tag{36}$$

assuming for now that the $l - 1$ subpopulations in the shortest path are not already diseased. Instead, in the case in which the re-routing mechanism to avoid infected subpopulations is active, individuals may go through alternative paths. Thus, we will have that:

$$D_{RR}^{new} = (l + \Delta l)(1 - R_0^{-n\Omega}) \tag{37}$$

where $\Omega$, $0 < \Omega \leq 1$ represents the ratio of infected individuals that have been rerouted. If all the $k - 1$ neighbors of the nodes before $r$ are not diseased we can evaluate

$$\Omega \sim \frac{1}{k-1} \tag{38}$$

Then, we take the ratio among the two quantities yielding

$$\frac{D_{RR}^{new}}{D^{new}} = \frac{(l + \Delta l)(1 - R_0^{-n\Omega})}{l(1 - R_0^{-n})}. \tag{39}$$

By expanding both the numerator and denominator for $R_0^{-n} \sim 1$, Eq. (39) can be written as

$$\frac{D_{RR}^{new}}{D^{new}} \sim (1 + \frac{\Delta l}{l})\Omega. \tag{40}$$

As $\Omega$ is of order $1/k$ (in the best case scenario one does not expect multiple change of paths), and $\Delta l$ and $l$ can be considered to be of the same order, we can expect

$$\frac{D_{RR}^{new}}{D^{new}} < 1, \tag{41}$$

The previous expression has been obtained for $R_0 \sim 1$ and $n$ relatively small. For large $R_0$ and $n$ values, we are in a different regime in which the probabilities $\left[1 - R_0^{-n}\right]$ 1 and $\left[1 - R_0^{-n\Omega}\right]$ 1 are independent of $\Omega$ thus defining

$$\frac{D_{RR}^{new}}{D^{new}} > 1. \tag{42}$$

Therefore, we expect a crossover from $D_{RR}^{new}/D^{new} < 1$ to $D_{RR}^{new}/D^{new} > 1$ in $R_0$ and $n$. As $n$ is an increasing function of the diffusion rate $\lambda$, the crossover is observed also for increasing values of this parameter. We plot the behavior of Eq. (39) in Fig. 4 for different values of $n$, $R_0$ and $\Omega$, setting $l = 3$ and $\Delta l = 1$ (plausible values in any random or complex networks) which confirms our findings from mechanistic numerical simulations discussed in the main text.

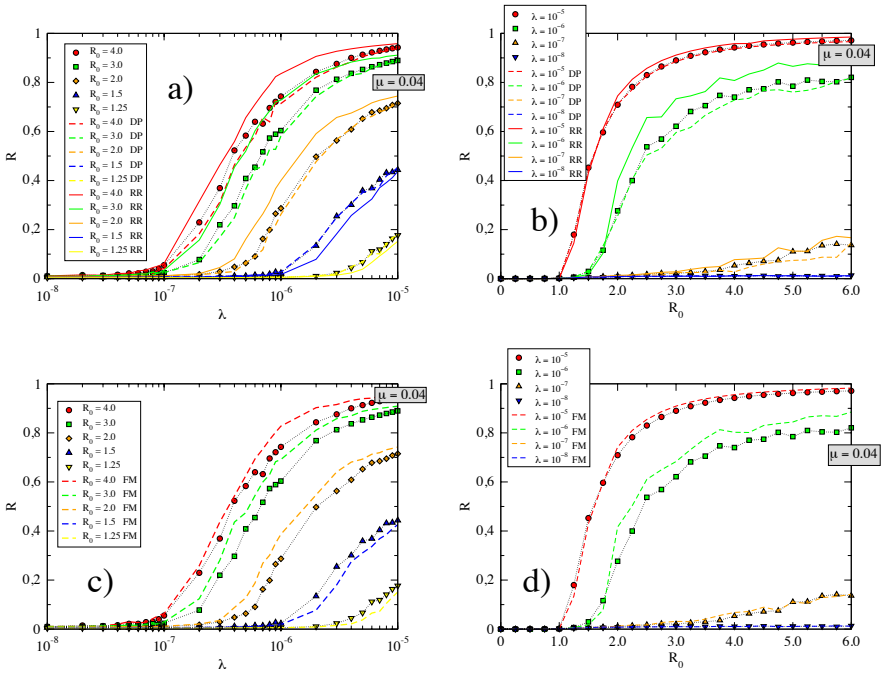


Figure 5: Comparison between the fraction of infected individuals $R$ obtained for the baseline model (i.e., when no behavioral changes are implemented) with that obtained with different risk aversion mechanisms as a function of $R_0$ or $\lambda$. Panels (a) and (b) show the results when the departure probability (DP) and the rerouting (RR) mechanisms are switched on separately. Panels (c) and (d) represent the same quantity but using the full model defined in the text. We have fixed $\mu = 0.04$. The underlying network is the worldwide air transportation network (see the text for further details).

### 3.4 Results

In this section we discuss the results obtained by implementing separately the two proposed mechanisms and when both are combined into a single, full model. We first notice that the critical invasion threshold $\lambda_c$ is the same for all the variants of the model. This critical point is independent of the behavioral changes because the model dynamics does not change at low values of the epidemic incidence, i.e., just around the global invasion threshold. In Fig. (5) we show the ratio of infected individuals $R$ (epidemic size) as a function of $\lambda$ (panel (a)) and $R_0$ (panel(b)). In the main text of the paper we show the same type of plots done instead for the ratio of diseased subpopulations $D/V$. The three different implementations of the model are shown as indicated in the legends: the baseline, the departure probability mechanism (DP) and the

rerouting mechanism (RR).

**Departure probability**

This mechanism brings a smaller reduction of the epidemic size for large values of $R_0$. In the other case instead the mechanism is not distinguishable from the baseline. This is due to the fact that for mild epidemics the number of infected individuals is small and then we can consider $\lambda_{ij} \sim \lambda$. Our results confirm previous findings about the inefficacy of measures like cutting down mobility of individuals by, for instance, closing airports [16, 17, 18, 19].

**Rerouting**

This mechanism brings, for values of $\lambda$ consistently bigger than its critical value, a bigger epidemic size with respect to the baseline model. This is due to the fact that, trying to minimize the risks to get sick, individuals change their route visiting places that otherwise would have not been visited, therefore contributing to a wider spread of the disease. As discussed in the preceding section, the previous behavior is not always obtained, since for low values of $R_0$ the number of individuals going through a node in the shortest path between an origin and a destination decreases due to the rerouting of individuals induced by the risk aversion mechanism. This is what happens, for instance, for $R_0 = 1.25$ and $R_0 = 1.5$ in Fig. (5)a and Fig. (5)b.

**Full model**

We implemented both mechanisms together in a full version of the model. As shown in Fig. (5)c and Fig. (5)d the effect of both risk aversion mechanisms on the epidemic size depends on the value of $R_0$ and $\lambda$. For $R_0 \geq 2$ and values of $\lambda$ larger than the global invasion threshold the full model leads to a bigger epidemic size when compared to the baseline implementation. In these regions of parameters, the rerouting mechanism is the dominant one and the large outbreak is due to the fact that people explore much more the network trying to minimize their risk to get sick. This self-fish strategy gives a worst scenario confirming how behavioral changes have a significant impact on the invasion dynamics in a non trivial way.

# 4 Model's Implementation

## 4.1 Internal nodes' dynamics

In each node a SIR dynamics takes place over a well mixed population of initial size $N_i(0) = w_i$, being $w_i$ the strength of node $i$. When time goes on, $N_i(t)$ changes according to the number of individuals that has been received and has left the node. Within the nodes, one step of a SIR process takes place. The

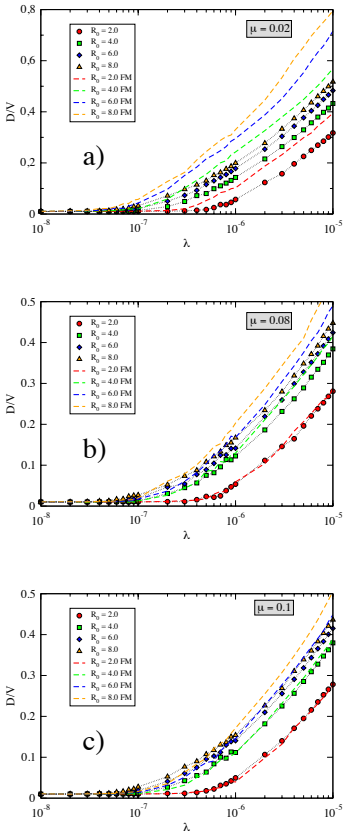


Figure 6: Comparison between the number of subpopulations affected by the outbreak, $D/V$, in the full ($\tau = 1$ and $\lambda^I = \lambda/2$) and null versions of the model for $\mu = 0.02, 0.08$ and $0.1$ as indicated. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.

state of every individual inside a node $i$ is changed according to the following probabilities: a susceptible individual becomes infected with probability $p^{(S\to I)} = 1 - (1 - \frac{\beta}{N_i})^{I_i}$, and an infected recovers with probability $p^{(I\to R)} = \mu$. Specifically, the exact number of individuals that changes its state is determined by a binomial distribution with the probability $p^{(S\to I)}$ (or $p^{(I\to R)}$) and the susceptible populations size $S_i(t)$ (or infected $I_i(t)$) as parameters. Note that in this scenario, $R_0$ only participates in the internal dynamics; individuals traveling through node $i$ are involved in the SIR and thus can change their state while at node $i$.

| Parameter | Range explored |
|---|---|
| $\mu$ | $0.02 - 0.1$ |
| $\tau$ | $\tau/3 - 3\tau$ |
| $\lambda^I$ | $\lambda/4 - \lambda$ |
| $h$ | 0.05−1.0 |
| $\gamma$ | $1/4 - 1$ |

Table 1: Summary of the sensitivity analysis performed.

### 4.2 Numerical simulations

The model is implemented as follows. Each simulation starts with a small number of infected individuals (1% of the individuals is infected within a randomly chosen subpopulation). In the simulations the traffic and spreading dynamics have the same time scale so, at each time step, first a diffusion step is performed and then the SIR internal dynamics is evaluated. For the diffusion of the individuals, we assume that the number of individuals starting a trip at time $t$ is given by a binomial distribution with the mobility rate $\lambda$ and the subpopulation size $N_i(t)$ as parameters. Destinations are chosen according to the population $w_i$ of each node. To keep the nodes' population constant over time (excluding transient individuals traveling through the node), once an individual reaches its destination it starts a new trip back to its origin. Simulations run until $I(t) = 0$ for all subpopulations. Note that in this framework only travelers are explicitly followed and that we only keep a counter for the number of non-traveling individuals inside each node.

## 5 Sensitivity Analysis

To test the sensitivity of our main results to different choices of the model parameters, we have performed further numerical simulations varying them. Namely, the dependency of the results presented throughout this work with the following parameters was explored: the recovery rate $\mu$, the time spent at destination before travelling back home $\tau$, the mobility rate of the individuals $\lambda^I$ and the re-routing intensity $h$. As an additional test we also extend our model including a further class of exposed individuals in which the transition from exposed to infected is governed by the transition rate $\gamma$. Table 1 summarizes the sensitivity results shown in the following subsections. Up to the range explored, the conclusions of our study remain the same. We note that given the large number of parameters involved and the large computational times required by the simulations, the sensitivity analysis has been restricted to values of the parameters within the typical ranges of epidemiological models or reasonable assumptions as given by the phenomenology of the models discussed.

## 5.1 Effects of recovery rate $\mu$

We have studied the model behavior for other three different values of $\mu$. In Fig. 6 we compare the results of the full model with $\tau = 1$ and $\lambda^I = \lambda/2$ for $\mu = 0.02$(a), $0.08$(b) and $0.1$(c) and four different values of $R_0 = 2, 4, 6, 8$. The results show that the behavior of the full model with respect to the baseline case is qualitatively the same regardless of the specific value of $\mu$.

## 5.2 Effects of the parameter $\tau$ accounting for the time spent at destination

In Fig. 7 we compare the results of the full model using three different values of $\tau$: $\tau' = \tau/2$ (a), $\tau' = 2 \cdot \tau$ (b) and $\tau' = 3 \cdot \tau$ (c), respectively, with the baseline case. No qualitative change of behavior is observed with respect to the results in the main text.

## 5.3 Effects of the mobility rate of infected individuals $\lambda^I$

Fig. 8 shows the results of the full model for three different values of the mobility rate of infected individuals $\lambda^I = \lambda/4$ (a),$\lambda^I = \lambda/3$ (b) and $\lambda^I = \lambda$ (c). Results corresponding to $\lambda^I = \lambda/2$ are presented in the main text. Note that although the value of $\lambda^I$ determines a variation on the global invasion threshold no qualitative change of behavior is observed with respect to the previous results. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.

## 5.4 Effects of rerouting intensity $h$

We have studied the model behavior for a family of $h$ values. As shown in Fig. 9, the full model always performs worse than the null case, provided that $h < 1/3$. When $h$ is above this value, all the curves collapse into a single family, whose behavior is nearly the one observed when individuals move following the shortest path.

## 5.5 Effects of latency rate $\gamma$ in SEIR model

Here we simulate an SEIR model assuming that individuals go from susceptible to exposed and then to the infected class at a rate $\gamma$. The latter is set to three different values $\gamma = 0.25$, $0.33$ and $0.5$ respectively. Fig. 10 shows that the addition of this new compartment does not alter the qualitative behavior of the model without such a class.

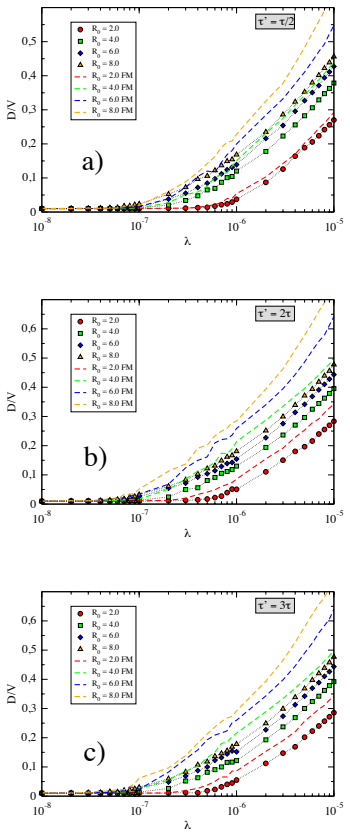


Figure 7: Comparison between the full and null versions of the model for $\mu = 0.04$ and three choices for the time spent at destination before traveling back home $\tau' = \tau/2, 2 \cdot \tau$ and $3 \cdot \tau$. $D/V$ is the number of subpopulations affected by the outbreak. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.

# References


[1] Colizza V. Vespignani V. (2008). Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations, *J. Theor. Biol.* 251:450-467

[2] Meloni, S., Arenas, A., Moreno, Y. (2009) Traffic-driven epidemic spreading in finite-size scale-free networks. *Proc. Nat. Acad. Sci. USA* 106:16897-16902.

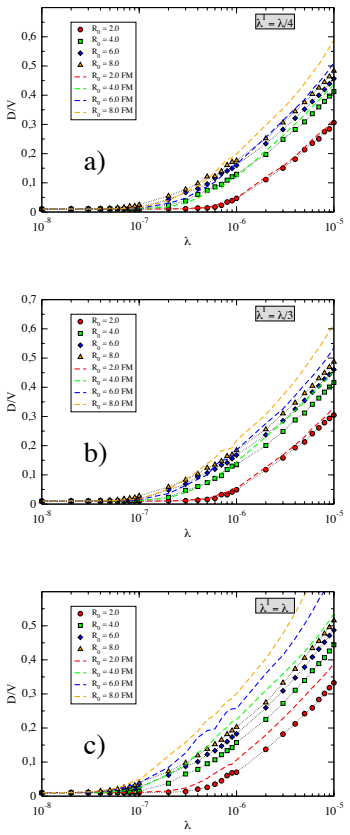


Figure 8: $D/V$ in the full and null versions of the model for $\mu = 0.04$ and mobility rates of infected individuals, $\lambda^I$, as indicated in the figure. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.


[3] Catanzaro, M., Boguña, M., Pastor-Satorras, R. (2005) Generation of uncorrelated random scale-free networks. *Phys. Rev. E* 71, 027103.

[4] Pastor-Satorras R., Vespignani A. (2001) Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* 86:3200-3203.

[5] Ferguson N.M. (2007). Capturing Human Behaviour, *Nature*, 466:733

[6] Brockmann D. Hufnagel L. Geisel L. (2006). The scaling laws of human travel, *Nature* 439:462-465.

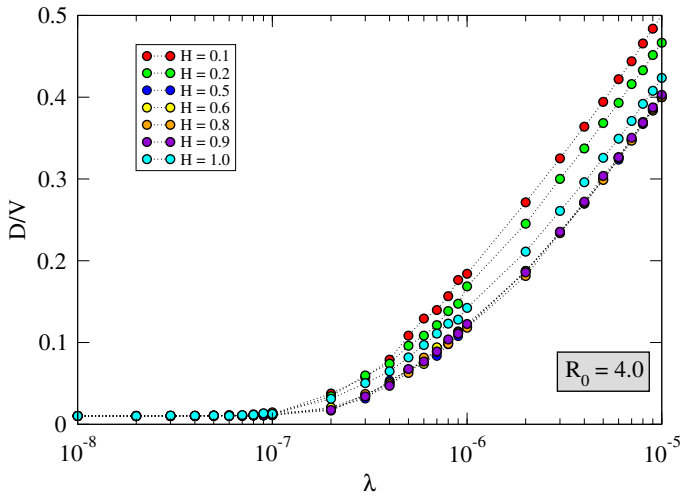


Figure 9: Results for $D/V$ obtained using different values of $h$ when the second risk aversion mechanism is active. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.

[7] Balcan D. Colizza V. Goncalves B. Hu H. Ramasco J.J. Vespignani A. (2009). Multiscale mobility networks and the large scale spreading of infectious diseases, *Proc. Natl. Acad. Sci. USA*, 106:21484.

[8] Hatchett R.J., Mecher C.E., Lipsitch M (2007) Public Health Interventions and Epidemic Intensity During the 1918 Influenza Pandemic, *Proc. Natl. Acad. Sci. USA* 104:7582-7587.

[9] Bootsma M.C.J., Ferguson N.M. (2007) The Effect of Public Health Measures on the 1918 Influenza Pandemic, *Proc. Natl. Acad. Sci. USA* 104: 7588-7593.

[10] Markel H., Lipman H.B., Navarro J.A., Sloan A., Michalsen J.R., Stern A.M., Cetron M.S. (2007) Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic, JAMA, 298:2260.

[11] Goffman W., Newill V.A. (1964) Generalization of Epidemic Theory an Application to the Transmission of Ideas, *Nature*, 204:225-228.

[12] Poletti P., Caprile B., Ajelli M., Pugliese A., Merler S. (2009). Spontaneous Behavioural Changes in Response to Epidemics, *J. Theor. Biol.* 241:193204.

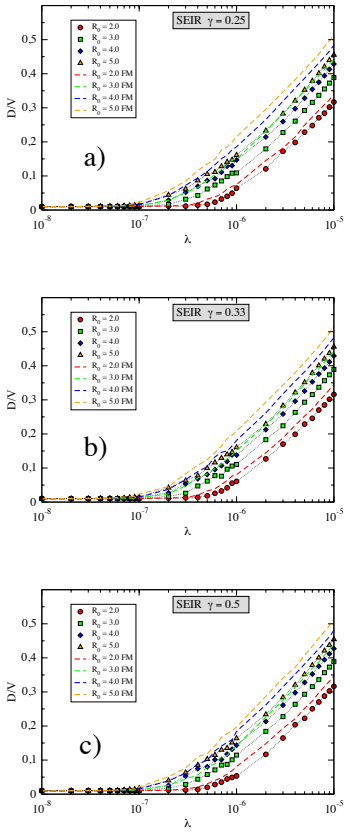


Figure 10: Comparison between the number of subpopulations affected by the outbreak, $D/V$, in the full and null versions of an extended model in which an exposed class has also been included. The transition from exposed to infected is set to $\gamma = 1/4$ (a), $\gamma = 1/3$ (b) and $\gamma = 1/2$ (c) , while the rest of parameters are $\mu = 0.04$, and $h = 0.1$. The network is the Air Transportation Network and destinations are chosen proportionally to nodes' strengths.

[13] Epstein, J.M., Parker, J., Cummings, D., Hammond R.A. (2008) Coupled Contagion Dynamics of Fear and Disease Mathematical and Computational Explorations. *PLoS ONE* 3(12):e3955.

[14] Funk, S., Gilad, E., Watkins, C., Jansen, V.A.A. (2009) The spread of awareness and its impact on epidemic outbreaks. *Proc. Nat. Acad. Sci. USA* 106:6872-6877.

[15] Mexican Authorities, *Boletín Mensual de Estadística Operacional. Secretaría*

*de comunicaciones y transportes (2009).*

[16] Colizza, V., Barrat, A., Barthelemy, M., Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Nat. Acad. Sci. USA* 103: 2015-2020.

[17] Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.J., Vespignani, A. (2007) Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions. *PLoS Med*, 4(1):e13.

[18] Germann, C., Kadau, K., Longini, I.M., Macken, C.A. (2006). Mitigation strategies for pandemic influenza in the United States. *Proc. Nat. Acad. Sci. USA* 103:5935-5940.

[19] Epstein, J.M. et al (2007) Controlling Pandemic Flu: The Value of International Air Travel Restrictions. *PLoS ONE* 2(5):e401.