



OPEN

## Endogenous social distancing and its underappreciated impact on the epidemic curve

Marko Gosak<sup>1,2</sup>, Moritz U. G. Kraemer<sup>3,4</sup>, Heinrich H. Nax<sup>5,6</sup>✉, Matjaž Perc<sup>1,7,8</sup> & Bary S. R. Pradelski<sup>9</sup>

Social distancing is an effective strategy to mitigate the impact of infectious diseases. If sick or healthy, or both, predominantly socially distance, the epidemic curve flattens. Contact reductions may occur for different reasons during a pandemic including health-related mobility loss (severity of symptoms), duty of care for a member of a high-risk group, and forced quarantine. Other decisions to reduce contacts are of a more voluntary nature. In particular, sick people reduce contacts consciously to avoid infecting others, and healthy individuals reduce contacts in order to stay healthy. We use game theory to formalize the interaction of voluntary social distancing in a partially infected population. This improves the behavioral micro-foundations of epidemiological models, and predicts differential social distancing rates dependent on health status. The model's key predictions in terms of comparative statics are derived, which concern changes and interactions between social distancing behaviors of sick and healthy. We fit the relevant parameters for endogenous social distancing to an epidemiological model with evidence from influenza waves to provide a benchmark for an epidemic curve with endogenous social distancing. Our results suggest that spreading similar in peak and case numbers to what partial immobilization of the population produces, yet quicker to pass, could occur endogenously. Going forward, eventual social distancing orders and lockdown policies should be benchmarked against more realistic epidemic models that take endogenous social distancing into account, rather than be driven by static, and therefore unrealistic, estimates for social mixing that intrinsically overestimate spreading.

The contact rates of infectious and non-infectious agents play a key role in determining the epidemic curve. How effective policies for (partially or fully) voluntary contact reductions—such as social distancing recommendations, or orders without monitoring or penalty—are depends on individual decisions in the population.

The current COVID-19 pandemic illustrates that governments disagree quite fundamentally regarding how much freedom of choice their citizens can or ought to be entrusted with to achieve desirable levels of contact reductions, resulting in less (e.g. Sweden<sup>1</sup>) and more (e.g. China<sup>2</sup>) stringent policies. To identify adequate policy responses for a given population, it is important to understand both, how it will react during a pandemic, and how it will react to policies being introduced, that is, to understand endogenous *behavioral change*<sup>3</sup>? in order to formulate relevant epidemiological models.

Unfortunately, the interactive nature of behavioral change related to contact reductions and social distancing within a population during an infectious disease outbreak (compared with vaccination<sup>4–7</sup>, antiviral prophylaxis<sup>8</sup>, and travel<sup>9</sup>) has not been explored in much detail, in particular not regarding the individual-level game-theoretic foundations of social distancing, and how these compare with real-world evidence. Progress in this direction ought to be made, because game-theoretic analyses have shown that interactions can crucially shape the epidemic curve<sup>10–13</sup>, and modeling increasingly rests on rich assumptions regarding how individual behavior changes dynamically with the disease outbreak.

<sup>1</sup>Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia. <sup>2</sup>Faculty of Medicine, University of Maribor, Taborska ulica 8, 2000 Maribor, Slovenia. <sup>3</sup>Department of Zoology, University of Oxford, Mansfield Road, Oxford OX1 3SZ, UK. <sup>4</sup>Harvard Medical School, 25 Shattuck St, Boston 02115, USA. <sup>5</sup>Behavioral Game Theory, ETH Zurich, Clausiusstrasse 37, 8092 Zurich, Switzerland. <sup>6</sup>Institute of Sociology, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland. <sup>7</sup>Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan. <sup>8</sup>Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria. <sup>9</sup>Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France. ✉email: heinrich.nax@uzh.ch

Making this step is particularly timely in light of the ongoing COVID-19 modeling, which has been highly policy-relevant. Currently, assumptions in these simulation models regarding behavioral change have to be made without adequate empirical foundations. The early simulations “driving the world’s response to COVID-19”<sup>14</sup>, for example, were based on static estimates of social mixing. These produced extreme scenarios in terms peak and case numbers of the outbreak, which were instrumental in governments’ justifications of policy introductions including of social distancing orders and lockdowns, in particular in the UK, but also across (most of) the globe. Recommending which policy options would be best suited to prevent these extreme scenarios, (many of the same) modelers again had to make relatively arbitrary behavioral assumptions regarding how social distancing policies would be adopted in terms of contact reductions in the population—as this example for the evaluation of a ‘voluntary home quarantine’ policy from<sup>15</sup> illustrates: “Following identification of a symptomatic case in the household, all household members remain at home for 14 days. Household contact rates double during this quarantine period, contacts in the community reduce by 75%. Assume 50% of household comply with the policy”.

The exact numbers that are being assumed for comparative policy evaluations can obviously matter crucially. Hence, modeling such<sup>15</sup>, which currently is extremely policy-relevant, must move from making arbitrary assumptions to theoretically and empirically validated ones as soon as possible. Here, by integrating behavioral micro-foundations we make such a step, and draw on game theory to embed interactive decision-making of social distancing in an epidemiological model.

Many countries during the current pandemic have decided to introduce temporal restrictions on movement, contacts and social interactions, because the general impression was that voluntary social distancing recommendations would not be sufficient<sup>16</sup>. As economic and social costs of the pandemic and of the measures accumulate, governments increasingly consider restricting contacts less and more efficiently, that is, locally<sup>17</sup> and/or dependent on health and risk status instead of locking down the whole population<sup>18</sup>. The issue is that research has not yet provided empirical benchmarks for endogenous contact rates in disease scenarios, so it is unclear how such policies can be evaluated scientifically: ideally a policy is benchmarked against a set of counterfactuals given the disease, not compared with what was before the disease. As such data is not available, we believe richer theory is needed, in particular game theory to provide a better rationale for making certain assumptions. We propose such a model, and to validate its basic logic compare its predictions with observations of contact rates during two influenza seasons in the United Kingdom where human contact was not affected by any noteworthy government policy. We view the observed effects as reasonable ‘lower bounds’ for counterfactual analyses in terms of what levels of social distancing ought to be expected from a more severe pandemic such as COVID-19 if no policies except educational ones were implemented. Our results suggest that levels of endogenous social distancing, as data suggests have been occurring during recent influenza seasons, might already flatten epidemic curves substantially. Policies need to be effective at least compared to such counterfactuals to warrant their introduction.

## Modeling infectious disease dynamics

The close monitoring and detailed modeling of infectious disease outbreaks has become an increasingly active research focus in epidemiology since the seminal works by Ref.<sup>19–23</sup>. The emergent body of research has substantially improved our understanding of infectious disease dynamics and how to control them (e.g. vaccination, quarantine, social distancing policies, etc.<sup>24–26</sup>), which together with the increasing availability of relevant data has allowed to apply some of these models to real-world epidemics.

One key modeling aspect concerns the transmission of infectious pathogens via individual contacts between infectious and susceptible individuals<sup>27,28</sup>, which have been shown to differ dependent on demographic factors such as age and sex<sup>29–31</sup>. While a lot of the prior work focuses on reconstructing the transmission trees of observed epidemics<sup>32</sup>, or on their final size and geographic spread<sup>33,34</sup>, less attention has been paid to the role that individual decision-making regarding social distancing—weighing the risks of infecting and being infected—plays in shaping behavioral contact patterns that underlie these dynamics.

Yet contact-seeking/avoiding decisions are key drivers of human disease dynamics. Descriptive analyses have revealed remarkable consistency at multiple temporal and spatial scales in the absence of exogenous factors<sup>35–39</sup>, and change as a result of disease severity<sup>40</sup>. One roadblock for making progress has been that relevant data are typically collected independently of health status. See<sup>41</sup> for one of the first surveys conducted via telephone during an influenza season. While this is currently changing with growing numbers of health-tracking applications, there is no robust data that has been made available as of now. As a consequence, there is little empirical evidence on how human contact rates change depending on health status and as a function of disease incidence overall. Some related empirical work has been done to capture behavioral change in response to environmental disaster<sup>42</sup>, and travel restrictions<sup>43,44</sup>, but not in response to disease outbreaks except for some recent work that we shall discuss separately in the concluding remarks. Precisely this kind of insight, however, would be important to understand contact rate decisions, because the incentives to practice social distancing or not are different for healthy and for sick people. To improve predictions concerning the dynamics of diseases at the population level, and to understand what kinds of policies are actually appropriate, the next step is uncovering the behavioral determinants of contact patterns to better justify necessary assumptions.

A first rational-choice foundation of individual contact-seeking/avoiding behavior in response to an infectious disease in epidemiological models is a framework proposed by Fenichel et al.<sup>45</sup>. Contact rates of the resulting epidemiological model are no longer exogenous variables, but instead are determined co-evolutionarily with the dynamics of the disease itself. Their framework presents an individual risk assessment, presuming that individuals’ propensities to socially distancing increase with intensity and awareness of the disease due to the increased risks for contracting the disease. Simulations show that incorporating this type of individual decision-making changes predictions concerning the epidemic curve: flatter curves are the result, particularly if sick individuals also reduce their contacts.

The decision-theoretic framework by Ref.<sup>45</sup> is a first step toward integrating human behavior into disease modeling, especially as regards understanding the role of infection fear. To improve the behavioral micro-foundations further, we proceed in two ways. First, the contact-reduction results are checked against some data on contact patterns during the 2012 and 2013 flu epidemics in the United Kingdom. Second, we account for the interactive nature of contact decisions by extending the underlying theoretical framework to a game-theoretic model. By using game theory we can model not just the trajectory of the disease as a function of the underlying contact data, but more generally endogenize contact patterns by an interactive decision model *and* as determined by the dynamics of the disease (e.g., incidence rates). Such a co-evolutionary view on contact rates and disease dynamics may substantially advance the resulting model's predictive potential.

The core argument of this paper is that risk assessments underlying contact decisions are interactive, which we model by formulating the “social distancing game”. This game theory model permits us to produce testable individual-level comparative statics regarding how individuals will react during the outbreak of a disease and in response to others' contact patterns (checking against data from two influenza seasons in the UK). Looking ahead, the advantage of our game-theoretic modeling approach is that it becomes feasible to identify tipping points in the underlying dynamics (as in<sup>46,47</sup>), whose transitions may be explosive and differ fundamentally for marginally different starting conditions as compared to those predicted by a non-game-theoretic model<sup>48</sup>. Future epidemic modeling should therefore consider game-theoretic modeling so as to leverage possible social dynamics of equilibrium transitions to policy advantage, as has been used in other policy domains<sup>49,50</sup>.

Two very recent concurrent papers make progress in this direction, and future work could merge our line of analyses with theirs. The first is<sup>51</sup> which considers a theoretical model with endogenous contact rates where the two types of agents, sick (and infectious but not yet symptomatic) and healthy, who choose contact rates are in the same information set. Infected individuals stay at home with probability one. Their model generates the same contact rates for both types, and does not make predictions regarding interaction effects of the two. Our data indicates that health status leads to different contact patterns, and that symptomatic individuals also vary contact rates as a function of incidence. This is also an important feature of our simulations. A very nice feature the model in<sup>51</sup>, which points towards interesting avenues for future work, is an explicit treatment of the path dependency of equilibrium. The second paper is a related theoretical framework by Ref.<sup>52</sup> who do not consider endogenous contact reductions by infected individuals at all because they have no private benefit from it. In that sense their model is more similar to<sup>45</sup> than ours, but adds a Nash equilibrium analysis to it. Again, our data indicates that infected and infectious individuals do also reduce contact rates with incidence levels, and that there are interactions between contact rates of sick and healthy individuals. Pro-social concerns for the health of others, not just concerns for one's own health, clearly play a very important role.

In sum, the ambition of this paper is to integrate behavioral responses from a game-theoretic framework into classical epidemiological models that accounts for health status and includes self-protective and pro-social concerns. By doing so, we propose a new model, spell out its behavioral predictions, in particular regarding differential rates of social distancing. We compare theoretical results with empirical observations from the 2012 and 2013 influenza epidemic in the United Kingdom, and discuss implications for policy recommendations in light of the simulated epidemic curves our model generates. We compare the endogenous curve with curves that would result from interventions such as immobilizing certain fractions of the population.

## Methods

**Contact rates.** The key ingredients, implicitly behavioral ones, that determine the dynamics of epidemiological models are ‘contact rates’ which govern the frequency and likelihood of human interactions and therefore transmissions: *Where do you go? Who do you see? How do you make contact?* At the individual level, a change in contact rates may occur for symptom-specific medical reasons after contracting a disease that leads to reduced mobility for example. Moreover, a person, whether infected or not, may consciously decide to social distance, that is, to reduce contacts in light of various evolving risks (i.e. of spreading the disease and/or of contracting the disease) during an outbreak. Instead of using the social distancing terminology, we referred to social distancing as ‘(partial) self-quarantine/isolation’ in earlier versions of the paper, but adopt this jargon in line with<sup>45</sup> as is becoming standard.

To understand the implications of these endogenous phenomena, we need a model for how and why behavioral change occurs during outbreaks of infectious diseases. To do so, we extend the existing decision-theoretic model of<sup>45</sup> to allow for interactive decisions and *strategic* considerations as the risks of contracting and transmitting a disease depend on one's own contact patterns as well as on everyone else's level of social distancing. Therefore, we model the individual decision as dependent on others' decisions, and we identify the rational-choice predictions for these decisions. By combining the human perspective on decision-making including considerations of risks and interactions in this way—using game theory—we obtain new and testable predictions for how human contact patterns and mobility decisions interact.

To illustrate the interactive nature of the proposed problem, consider the following thought experiments at the two extremes of the logical spectrum. At one extreme, suppose that everyone (sick and healthy alike) stays home, i.e. has reduced their contacts to zero (extreme social distancing). In that case, of course, any given individual (think of Will Smith as the only daytime person on the streets of NYC in “I Am Legend” to lighten the mood) can move freely without fear of infection (if healthy) or of infecting others (if sick). Thus, in game-theoretic language, this does not constitute a Nash equilibrium, because every individual prefers to deviate (from staying at home), given the decision of everyone else (to stay at home). At the other extreme, by contrast, when everyone (sick and healthy) is moving around all the time resulting in very high contact rates (no social distancing), it is safest to stay home in order to not become infected (if healthy) or not to infect others (if sick). Again, everyone moving freely around will not constitute an equilibrium.

**Social distancing in a population.** In this section, we propose a formal model that will highlight the main advantages of choosing a game-theoretic rather than mechanistic approaches (as is done in applied work), and spell out how it goes beyond a single-player decision-theoretic model.

*Population.* Consider a human population  $N = \{1, 2, \dots, n\}$ . Each person  $i \in N$  either belongs to the set  $H \subset N$ , the *healthy* (or non-symptomatic, susceptible, uninfected, etc.), or to the other set  $S = N \setminus H$ , the *sick* (or symptomatic, non-susceptible, infected, etc.). Note that we work with a basic epidemic model setting without recovery (and repeat susceptibility) in mind, which naturally ought to be generalized in future work.

*Social-distancing decisions.* Each  $i \in N$  chooses a contact rate  $\beta_i \in [0, 1]$ . Write  $\beta$  for the full vector of contact rates,  $\beta_H$  for the average contact rate of healthy agents, and  $\beta_S$  for the average contact rate of sick agents.

*Utilities.* Individual utility is generated by reaching places (or people) which is facilitated by being mobile. Hence, positive mobility is required to generate utility. But increased levels of mobility are also increasingly costly as they increase the exposure to infection risks for self and others. Hence, both complete immobility and full mobility generate no utility. Once there are risks of infection due to the presence of a disease, this mobility will be reduced to mitigate these risks.

Let us consider two scenarios distinguished by whether (1) everyone is healthy, or (2) there are infected individuals.

(1) *No-disease scenario* Suppose there is no disease, that is,  $|S| = 0$ . In that case, we assume utility for any player  $i$  is described by a twice-differentiable, continuous utility function

$$\text{Base utility. } u_i(\beta) = u(\beta_i) \quad (1)$$

such that  $u(0) = u(1) = 0, u(\beta) > 0$  for all  $\beta \in (0, 1), u'(\beta) > 0 (< 0)$  for  $\beta < \beta^* (> \beta^*)$  given some  $\beta^* \in (0, 1)$ , and  $u''(\beta) < 0$ . These assumptions ensure that  $\beta^*$  represents the unique utility-maximizing level of mobility in the no-disease scenario. We shall refer to levels chosen below  $\beta^*$  as ‘social distancing’. Of course, the optimal level will be heterogeneous within a population, but we abstract from this level of detail for the moment. Recent empirical work by Ref.<sup>38,50</sup> identifies heterogeneous levels of mobility in the absence of a disease.

(2) *Disease scenario* Once some individuals are infected, that is, if  $|S| > 0$ , then the ‘base utility’  $u(\beta_i)$  that corresponds to a healthy individual, for an infected individual, is reduced directly by some disease factor  $\delta$  (with  $\delta \in [0, 1]$ ) representing a proportional disutility from being sick) resulting in ‘sick utility’  $\delta u(\beta_i)$ . Moreover, depending on health status, all individuals suffer additional disutility from the risk of becoming infected (for healthy), or from the risk of infecting others (for sick), both of which increase with mobility, thus adding further costs to being mobile. Hence, for a **healthy individual**,  $i \in H$  the utility is

$$\text{Healthy } H\text{-utility } u_i(\beta) = \underbrace{(1 - f \cdot [(1 - \beta_S^i)^{n-|H|} \cdot \beta_i + 1 - \beta_i])}_{\text{infection risk}} \cdot \underbrace{u(\beta_i)}_{\text{base utility}} \quad (2)$$

where  $f \in [0, 1]$  measures the fear of a healthy individual of getting infected, which also expresses disease severity. Similarly,

$$\text{Sick } S\text{-utility. } u_i(\beta) = \underbrace{(1 - c \cdot [(1 - \beta_H^i)^{|H|} \cdot \beta_i + 1 - \beta_i])}_{\text{spreading risk}} \cdot \underbrace{\delta u(\beta_i)}_{\text{sick utility}} \quad (3)$$

where  $c \in [0, 1]$  measures the pro-social concern an infected individual has for another individual’s life, that is, the expected reduction in utility from exposing other healthy humans to the risk of infection, which would naturally increase with the severity of the disease too. Note that the introduction of this parameter expressing this type of motivation, which is central to most policies aimed at reducing mobility of symptomatic humans, is absent in<sup>45</sup>, but will generate the kinds of mobility reductions that characterize several of his simulations resulting in the flattest epidemic curves.

The underlying contact scenario we thus express is one where  $\beta_i$  represents agent  $i$ ’s probability of exposing him/herself to an infection-risk encounter, and  $1 - (1 - \beta_S^i)^{n-|H|}$  and  $1 - (1 - \beta_H^i)^{|H|}$  respectively represent the probabilities of at least one infected/susceptible making the same encounter. Thus we model the probability of two parties meeting at a given location, or all parties spending some time at a central locations. W.l.o.g., when two individuals with different health status enter the location, we assume an infection takes place with probability one.

**Simulation details.** We use random geometric graphs in hyperbolic spaces to generate networks that have heterogeneous degree distributions, strong clustering, and short average path lengths, which are all inherent properties of real social networks<sup>53,54</sup>. By increasing the curvature  $\zeta$  of the hyperbolic space, we move from networks having exponential to networks having scale-free degree distributions, from longer to shorter average path lengths, and from weaker to stronger clustering. We thus cover the whole family of networks that are representative for real social networks<sup>55</sup>.

On top of these networks, we consider the susceptible-exposed-infectious-recovered (SEIR) model<sup>56,57</sup>, as used for describing the spreading of the COVID-19 disease<sup>58</sup>. Initially, we select 0.2% of the nodes uniformly at random and designate them as infected (I). The remaining 99.8% of the nodes are designated as susceptible (S). Moreover, every node  $i$  is assigned a contact rate  $q_i$ , where  $q_i = 0$  means the node is not exposed at all and thus has no way of becoming infected, while  $q_i = 1$  means the node is fully exposed to potentially become infected

A marginal increase in ...	... leads to ...
Social distancing of healthy	Less social distancing of sick
Social distancing of sick	Less social distancing of healthy
Size of the healthy population	More social distancing of sick
Size of the sick population	More social distancing of healthy
Pro-social concern of the sick	More social distancing of sick
Dear of disease of the healthy	More social distancing of healthy*

**Table 1.** Comparative statics of the equilibrium analysis. These describe how optimal contact rates vary with the other parameters, evaluated under the assumption that a symmetric Nash equilibrium exists such that  $\beta_i = \beta_H$  for all healthy and  $\beta_i = \beta_S$  for all sick. See Supplementary Information ‘Comparative Statics Derivations’ for details. \*Present in the model by Ref.<sup>45</sup>. The other effects are new ‘Contrary to imitation, herding, etc. as proposed, for example, in<sup>11,59</sup>’.

by all the other nodes to which it is connected. In terms of social distancing, we can think of  $q_i$  as the factor multiplied to the optimal baseline level of contact-making behavior  $\beta^*$  from our model.  $q_i = 0$  means that node  $i$  is fully isolated, while  $q_i = 0.5$  means there is a 50% chance node  $i$  will make contact with any other nodes to which it is connected, and  $q_i = 1$  means full contact-making behavior without social distancing as given by  $\beta^*$ . We consider several models; the model without social distancing with  $q_i = 1$  for all nodes, the model with uniform social distancing such that  $q_i < 1$  for all nodes, the model with random social distancing such that a fraction  $p$  of nodes is selected at random and assigned  $q_i = 0.1$  instead of  $q_i = 1$ , and, finally, the model motivated by our game-theoretic analysis with endogenous social distancing fitted against the Flusurvey data to account for decreasing  $q_i$  as the fraction of infected individuals  $\rho$  in the population increases. In particular, the function we use is  $q_i = 3^{(-10\rho)}$ , which yields a three-fold decrease in  $q_i$  at 10% of infected in the population (which we approximate by extrapolating from the behavior of ill individuals from the observed values between 3 and 8% of infected in the population where we fitted such a slope—see top left of Fig. 3). We refer to the Supplementary Information files for details concerning the Monte Carlo method that we use to simulate this model including the code.

### Comparative statics

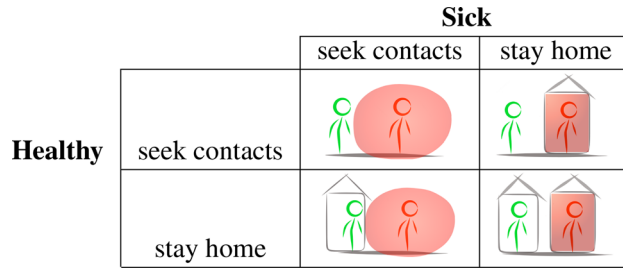
What interests us from the game-theoretic model are the comparative statics of contact rates in equilibrium when chosen optimally so as to maximize subjective expected utilities. These we obtain from the first-order conditions for optimal behavior for the two utility functions given by Eqs. (2) and (3), which we obtain by maximizing both expressions with respect to  $\beta_i$ :

$$\begin{aligned}
 H\text{-FOC. } & \overbrace{(1 - f \cdot (1 - [(1 - \beta_S^i)^{n-|H|} \cdot \beta_i + 1 - \beta_i]))u'(\beta_i)}^{\text{marginal utility effect}} \\
 & = \underbrace{f \cdot (1 - (1 - \beta_S^i)^{n-|H|})u(\beta_i)}_{\text{marginal infection risk effect}}
 \end{aligned}
 \tag{4}$$

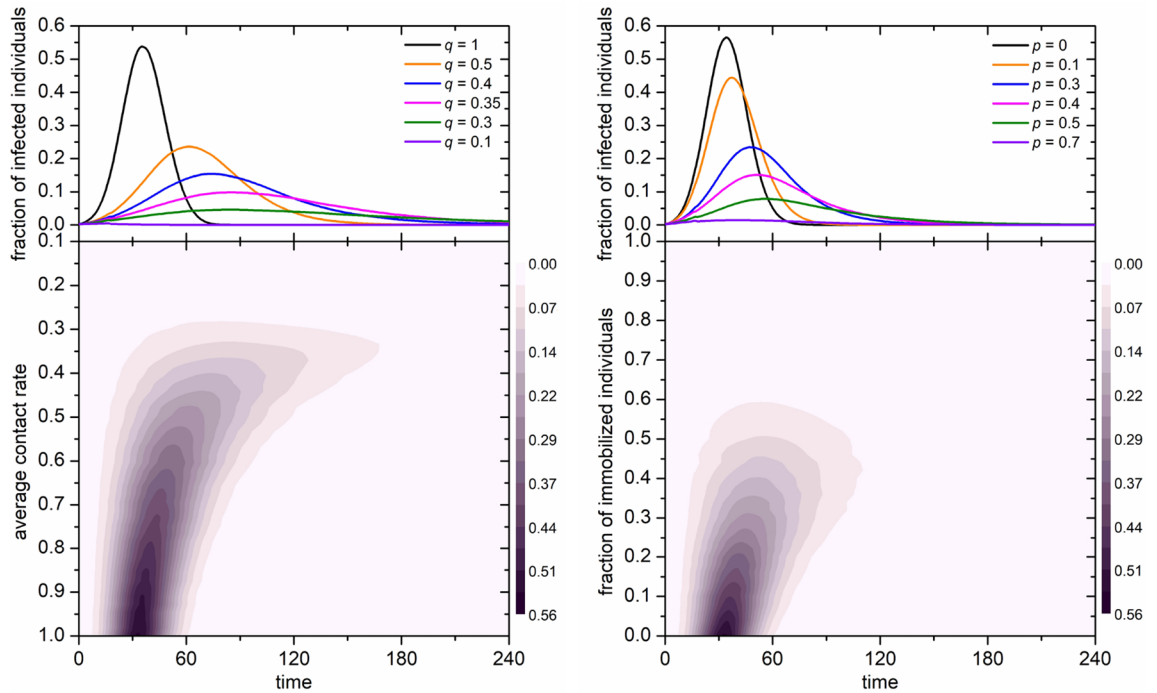
$$\begin{aligned}
 S\text{-FOC. } & \overbrace{(1 - c \cdot (1 - [(1 - \beta_H^i)^{|H|} \cdot \beta_i + 1 - \beta_i]))u'(\beta_i)}^{\text{marginal utility effect}} \\
 & = \underbrace{c \cdot (1 - (1 - \beta_H^i)^{|H|})u(\beta_i)}_{\text{marginal spreading risk effect}}
 \end{aligned}
 \tag{5}$$

Note that both right-hand sides of the latter equations are positive, indicating that both marginal utilities  $u'(\beta_i)$ s must also be positive; i.e. that we now must obtain lower contact rates for both sick and healthy individuals (compared with the utility-maximizing level of mobility  $\beta^*$  from the no-disease scenario) in order for FOCs to be satisfied than in the no-disease benchmark. This means that both sick and health individuals will engage in some optimal level of ‘social distancing’, that is, choosing a lower equilibrium utility than  $\beta^*$  from the no-disease scenario. The comparative statics are summarized in Table 1.

Naturally, the optimal contact rates for healthy and sick are different and take intermediate values, the exact value depending on factors related to disease incidence, fear, concern, disease severity, risks, etc. Note that individuals may also differ in their fears, concerns, etc., hence we can think of the comparative statics in Table 1 also as organizing individual heterogeneity. While this is a two-population evolutionary game with continuous action space for every player, the strategic essence of this interaction can be represented by a simplified game played between Sick and Healthy as is illustrated in Fig. 1. Both health types seek contact leads to infection. Both staying home leads to no infection, but also generates zero utility for anyone. The two mixed outcomes, where only one party stays home, also do not lead to infection, and have the advantage that the population that continues to be mobile generates positive utilities. Similarly, if the elderly are particularly at risk, either the young or the elderly,



**Figure 1.** The social distancing interactions simplified.



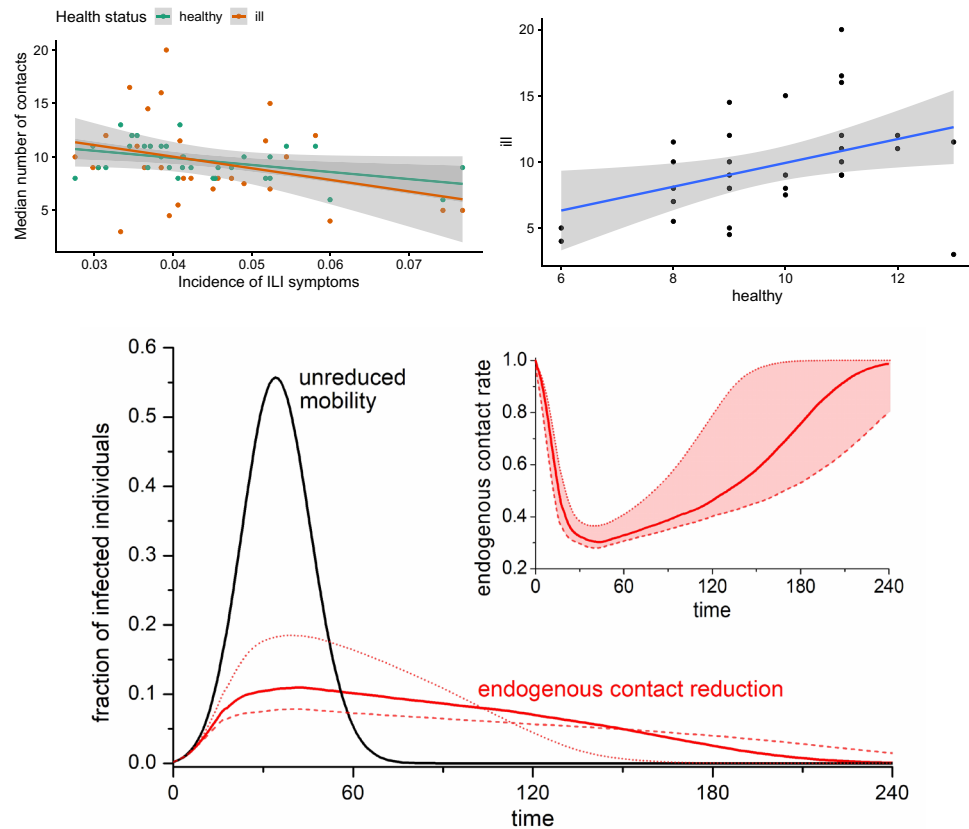
**Figure 2.** Simulated epidemic curves based on a susceptible-exposed-infectious-recovered (SEIR) model obtained by averaging outcomes over 1.5 million nodes in network configurations representing social networks (see the Simulation details section above and the Monte Carlo method section in Supplementary Information). Left: The color map encodes the fraction of infected individuals in dependence on time and the average contact rate. The upper panel shows characteristic cross-sections of the color map, where it can be observed that the endogenous contact reduction effect is matched no sooner than at 60–70% reduction of the contact rate (average mobility). Right: The color map encodes the fraction of infected individuals in dependence on time and the fraction of immobilized individuals. The upper panel shows characteristic cross-sections of the color map, where it can be observed that the endogenous contact reduction effect is matched no sooner than at 40–50% immobilization.

or both, should perhaps avoid contacts with one another to avoid infections. As governments aim to return to higher levels of economic and social activities, such an outcome, with the sick (the high-risk groups) rather than the healthy rest reducing contacts, will likely become the goal.

Instead of assuming a particular functional form to obtain contact rates, we shall explore whether the direction of contact rate adjustments as per our comparative statics corresponds qualitatively with data from the UK Flusurvey (see Fig. 3). Flusurvey is a webtool managed and monitored by Public Health England (PHE). All experimental protocols were approved by the London School of Hygiene and Tropical Medicine Ethics Committee, and were carried out in accordance with relevant guidelines and regulations. Participation is voluntary, and informed consent was obtained from all subjects.

### Concluding remarks

*Who makes which contacts when?* is what needs to be modeled to obtain a realistic epidemic curve. An under-appreciated element of contact behaviors is their interactive nature, and the cost-benefit analyses driving them. Health status, infection risk and behaviors of others matter for the individual decision in ways that



**Figure 3.** Social distancing for sick and healthy, and the effect on the epidemic curve. Top left: Median number of contacts in different weeks as a function of incidence of ILI symptoms among Flusurvey participants during that week. The lines show linear fits, and shades 95% confidence intervals. Slopes: healthy – 70 (95% CI – 120–(– 20)), ill – 110 (95% CI – 220, 10);  $p$  value testing null hypothesis of slope 0: healthy 0.01, ill 0.07. Top right: Median number of contacts in participants with ILI symptoms as a function of the median number of contacts in participants without ILI symptoms. The line shows a linear fit, and shades 95% confidence intervals. Slope: 0.9 (95% CI 0.2–1.6);  $p$  value testing null hypothesis of slope 0: 0.02. Bottom: Comparison of the infected curves under unrestrained mobility (black) and endogenous contact reduction based on the Flusurvey data (black). The inset shows how fast the contact rate decreases as the fraction of infected individuals peaks, and then increases comparatively slowly as the incidence of infections decreases. Dashed and dotted lines were obtained with three-fold reductions at 5% and 15%, respectively, as the lower and upper bound on the error from the data (which suggest a three-fold decrease at about 9%/15% for ill/healthy—compare with slopes – 110/– 70 from top left). For a description of the model and simulations, see the Simulation details section above and the Monte Carlo method section in Supplementary Information.

state-of-the-art epidemiological modeling does not capture. In this paper, we proposed a rational-choice framework to endogenize these decisions and the resulting contact rates. While promising conceptually, a major issue for modeling behavior in general, and to test the kinds of predictions that our model generates in particular, is data availability. To date, very little data is available that records contact rates and health status at the same time. This might soon change as health-tracking applications are increasingly adopted.

As a first step, we simulated different interpretations of social distancing policies in Fig. 2, highlighting what kinds of epidemic curves ought to be expected from either reducing mobility of everyone in the population or from immobilizing a certain fraction. Towards some empirical foundations of these kinds of analyses, we considered Flusurvey data from the United Kingdom (see Fig. 3, and Supplementary Information ‘Influenza Contact Data’), where we found evidence of social distancing amongst healthy individuals as a function of disease incidences in their neighborhoods, as was predicted by our model and by earlier work. We pre-registered the basic notion of this kind of empirical analysis qualitatively before knowing what kind of data would be available exactly (see Open Science Framework projects [osf.io/zc5b8](https://osf.io/zc5b8) and [osf.io/q3m2p](https://osf.io/q3m2p)). The baseline levels of contact making in the UK as recorded per Flusurvey (resulting in medians of circa 12–14 contacts per week outside the flu season) are in line with prior estimates from other countries than the UK<sup>41,60</sup>. Indeed, we found that, even in the context of seasonal influenza, some sizeable degree of social distancing took place amongst both sick (mobility reduction of ca. 50–55%) and healthy individuals (mobility reduction of ca. 30–35%), in line with endogenous social distancing at the population level in other countries<sup>61</sup>. We extrapolated these kinds of reductions linearly to levels outside the observed range in our simulations. Predicted negative correlations between the levels of the

two health types were rejected, suggesting presence of behavioral elements beyond individual utility maximization such as social influence, norms, imitation, herding, etc.

We consider the fitting with influenza data useful, because a seasonal influenza virus is less severe than other pandemics such as SARS-CoV-2, so any endogenous mobility reductions observed for influenza would likely provide lower bounds on the reductions that we would expect (without policy) during a more severe pandemic such as the current one. In Sweden, for example, where the government decided against the kinds of lockdowns that other European countries implemented, the aggregate population mobility in transit and workplace decreased by 31% and 11% respectively between onset and peak of the pandemic (as per Google's COVID-19 Community Mobility Report Sweden), which is comparable to the decrease we recorded in the Flusurvey. Our simulations in Fig. 3, which use the fit for contact reductions as observed in Flusurvey—extrapolating further reductions linearly in case of incidence levels beyond those observed, indicate that such reductions would flatten the epidemic curve to levels that are comparable in terms of height of the peak and total case numbers as would have been obtained from immobilizing 40–50% of the total population or bringing the average mobility down by 60–70%. These are candidate benchmarks we should be evaluating policy success against, not against historical data.

We are hopeful that future research and applied modeling will make use of game theory with the modeling framework we proposed. Further, we encourage future efforts to test our model's hypotheses with more data and with data for COVID-19 instead of influenza, as there are potential confounding factors in our data related to the seasonality of contacts because of factors unrelated to disease (especially temperature, but also school holidays, etc.), which we cannot account for sufficiently due to data availability. Such analyses are important, as policymakers consider new, perhaps health-status dependent, mobility restrictions and relaxations thereof during the current pandemic. Ideally, to evaluate the effectiveness of such policies there ought to be at least some benchmarking concerning what levels might be expected endogenously, as well as monitoring of individual behaviors in response to changes in government recommendations or restrictions. For example, the UK COVID-19 lockdown from earlier this year is estimated to have reduced contacts by 75%<sup>62</sup>, which is roughly double the reduction we recorded for healthy individuals for the 2012 and 2013 influenza seasons (see Fig. 3), but not substantially above the levels that some of our simulations indicate would justify such policies (see Fig. 2). This work suggests scope for future studies in this directions and provides some first measurements.

Governments should factor in endogenous social distancing when weighing the pros and cons of policies as diverse as those ranging from China to Sweden. Epidemic modeling could improve its behavioral micro-foundations more generally.

## Data availability

The data underlying the analyses of this article is available upon request from Flusurvey.

Received: 25 August 2020; Accepted: 21 January 2021

Published online: 04 February 2021

## References

- Lindström, M. The COVID-19 pandemic and the Swedish strategy: epidemiology and postmodernism. *SSM Popul. Health* **11**, 100643 (2020).
- Kraemer, M. U. *et al.* The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
- Manfredi, P. & D'Onofrio, A. *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (Springer, Berlin, 2013).
- Bauch, C. T., Galvani, A. P. & Earn, D. J. Group interest versus self-interest in smallpox vaccination policy. *Proc. Natl. Acad. Sci.* **100**, 10564–10567 (2003).
- Bauch, C. T. & Earn, D. J. Vaccination and the theory of games. *Proc. Natl. Acad. Sci.* **101**, 13391–13394 (2004).
- Galvani, A. P., Reluga, T. C. & Chapman, G. B. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. *Proc. Natl. Acad. Sci.* **104**, 5692–5697 (2007).
- Fu, F., Christakis, N. A. & Fowler, J. H. Dueling biological and social contagions. *Sci. Rep.* **7**, 43634 (2017).
- van Boven, M., Klinckenberg, D., Pen, I., Weissing, F. J. & Heesterbeek, H. Self-interest versus group-interest in antiviral control. *PLoS ONE* **3**, e1558 (2008).
- Zhao, S., Bauch, C. T. & He, D. Strategic decision making about travel during disease outbreaks: a game theoretical approach. *J. R. Soc. Interface* **15**, 20180515 (2018).
- Reluga, T. C. An SIS epidemiology game with two subpopulations. *J. Biol. Dyn.* **3**, 515–531 (2009).
- Poletti, P., Caprile, B., Ajelli, M., Pugliese, A. & Merler, S. Spontaneous behavioural changes in response to epidemics. *J. Theor. Biol.* **260**, 31–40 (2009).
- Reluga, T. C. Game theory of social distancing in response to an epidemic. *PLoS Comput. Biol.* **6**, e1000793 (2010).
- Glaubit, A. & Fu, F. Oscillatory dynamics in the dilemma of social distancing. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **476**, 20200686. <https://doi.org/10.1098/rspa.2020.0686> (2020).
- Adam, D. Special report: the simulations driving the world's response to COVID-19. *Nature* **580**, 316 (2020).
- Ferguson, N. M. *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce covid-19 mortality and healthcare demand. *DOI* **10**, 77482 (2020).
- Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
- Pradel, B. & Oliu-Barton, M. 11 Green bridges: reconnecting Europe to avoid economic disaster. *Eur. Time Covid-19* **83** (2020).
- Ferretti, L. *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, 6936 (2020).
- Ross, R. An application of the theory of probabilities to the study of a priori pathometry. Part I. *Proc. R. Soc. A* **92111184**, 204–230 (1916).
- Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry. Part III. *Math. Phys. Charact.* **93**, 225–240 (1917).
- Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry. Part II. *Math. Phys. Charact.* **93**, 212–225 (1917).



22. Kermack, W. O. & McKendrick, A. G. Mathematical theory of epidemics. Part II. The problem of endemicity. *Proc. R. Soc. A* **115**, 55–73 (1927).
23. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **115**, 700–721 (1927).
24. Anderson, R. M. & May, R. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1991).
25. Heesterbeek, H. *et al.* Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, aaa4339 (2015).
26. Litvinova, M., Liu, Q.-H., Kulikov, E. S. & Ajelli, M. Reactive school closure weakens the network of social interactions and reduces the spread of influenza. *Proc. Natl. Acad. Sci.* **116**, 13174–13181 (2019).
27. Cauchemez, S. *et al.* Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N. Engl. J. Med.* **361**, 2619–27 (2009).
28. Read, J. M., Edmunds, W. J., Riley, S., Lessler, J. & Cummings, D. A. Close encounters of the infectious kind: methods to measure social mixing behaviour. *Epidemiol. Infect.* **140**, 2117–2130 (2012).
29. Wallinga, J., Teunis, P. & Kretzschmar, M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* **164**, 936–44 (2006).
30. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).
31. Cauchemez, S. *et al.* Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proc. Natl. Acad. Sci.* **108**, 2825–30 (2011).
32. Cauchemez, S., Carrat, F., Viboud, C., Valleron, A. J. & Boëlle, P. Y. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat. Med.* **23**, 3469–3487 (2004).
33. Viboud, C. *et al.* Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–51 (2006).
34. Kraemer, M. U. G. *et al.* Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. *Lancet Infect. Dis.* **17**, 330–338 (2017).
35. González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding individual human mobility patterns. *Nature* **453**, 779–82 (2008).
36. Candia, J. *et al.* Uncovering individual and collective human dynamics from mobile phone records. *J. Phys. A Math. Theor.* **41**, 224015 (2008).
37. Simini, F., González, M. C., Maritan, A. & Barabási, A. L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
38. Alessandretti, L., Sapiezynski, P., Sekara, V., Lehmann, S. & Baronchelli, A. Evidence for a conserved quantity in human mobility. *Nat. Hum. Behav.* **2**, 485 (2018).
39. Luh, D.-L., You, Z.-S. & Chen, S.-C. Comparison of the social contact patterns among school-age children in specific seasons, locations, and times. *Epidemics* **14**, 36–44 (2016).
40. Kerckhove, K. V., Hens, N., Edmunds, W. J. & Eames, K. T. D. The impact of illness on social networks: implications for transmission and control of influenza. *Am. J. Epidemiol.* **178**, 1655–1662 (2013).
41. DeStefano, F. *et al.* Factors associated with social contacts in four communities during the 2007–2008 influenza season. *Epidemiol. Infect.* **139**, 1181–1190 (2011).
42. Lu, X., Bengtsson, L. & Holme, P. Predictability of population displacement after the 2010 Haiti earthquake. *Proc. Natl. Acad. Sci.* **109**, 11576–11581 (2012).
43. Peak, C. M. *et al.* Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. *Int. J. Epidemiol.* **47**, 1562–1570 (2018).
44. Lam, E. H. Y. *et al.* The feasibility of age-specific travel restrictions during influenza pandemics. *Theor. Biol. Med. Model.* **8**, 1–14 (2011).
45. Fenichel, E. P. *et al.* Adaptive human behavior in epidemiological models. *Proc. Natl. Acad. Sci.* **108**, 6306–6311 (2011).
46. Centola, D., Becker, J., Brackbill, D. & Baronchelli, A. Experimental evidence for tipping points in social convention. *Science* **360**, 1116–1119 (2018).
47. Centola, D. & Baronchelli, A. The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proc. Natl. Acad. Sci.* **112**, 1989–1994 (2015).
48. Young, H. P. The dynamics of social innovation. *Proc. Natl. Acad. Sci.* **108**, 21285–21291 (2011).
49. Nyborg, K. *et al.* Social norms as solutions. *Science* **354**, 42–43 (2016).
50. Hu, H., Nigmatulina, K. & Eckhoff, P. The scaling of contact rates with population density for the infectious disease models. *Math. Biosci.* **244**, 125–134 (2013).
51. McAdams, D. Nash SIR: an economic-epidemiological model of strategic behavior during a viral epidemic. *Covid Econ.* **115–135**, 20 (2020).
52. Toxvaerd, F. Equilibrium social distancing. *Covid Econ.* 110–134.
53. Boguná, M., Papadopoulos, F. & Krioukov, D. Sustaining the internet with hyperbolic mapping. *Nat. Commun.* **1**, 62 (2010).
54. Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. & Boguná, M. Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, 036106 (2010).
55. Zuev, K., Boguná, M., Bianconi, G. & Krioukov, D. Emergence of soft communities from geometric preferential attachment. *Sci. Rep.* **5**, 9421 (2015).
56. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015).
57. Wang, Z. *et al.* Statistical physics of vaccination. *Phys. Rep.* **664**, 1–113 (2016).
58. Prem, K. *et al.* The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* **5**, e261–e270 (2020).
59. Poletti, P., Ajelli, M. & Merler, S. Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. *Math. Biosci.* **238**, 80–89 (2012).
60. Kretzschmar, M. & Mikolajczyk, R. T. Contact profiles in eight European countries and implications for modelling the spread of airborne infectious diseases. *PLoS ONE* **4**, e5931 (2009).
61. Poletti, P., Ajelli, M. & Merler, S. The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics. *PLoS ONE* **6**, e16460 (2011).
62. Jarvis, C. I. *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med.* **18**, 1–10 (2020).

## Acknowledgements

We are thankful for comments from several colleagues, in particular from Petra Klepac and Bernhard von Stengel. We thank Sebastian Funk for access to Flusurvey data. M.G. acknowledges funding from the Slovenian Research Agency (Grant No. P3-0396). M.U.G.K. acknowledges funding from the Branco Weiss Fellowship, Oxford Martin School and European Horizon 2020 Programme project MOOD. H.H.N. acknowledges an SNF Eccellenza Grant from the Swiss National Science Foundation. M.P. acknowledges funding from the Slovenian Research Agency (Grant Nos. J1-2457, J1-9112, and P1-0403).

### Author contributions

All authors contributed to all aspects of the paper, with M.U.G.K. and H.H.N. designing the study, and B.S.R.P. and H.H.N. modeling, M.G. and M.P. simulating based on Flusurvey.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-82770-8>.

**Correspondence** and requests for materials should be addressed to H.H.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2021