Supplementary Information

To isolate, or not to isolate: a theoretical framework for disease control via contact tracing

Davin Lunz, Gregory Batt, Jakob Ruess

S1 Modeling framework

S1 A Disease model with testing and contact tracing

In this section we pose the disease dynamics and testing structure, and derive the contact tracing contribution from first principles.

Disease dynamics

We consider a compartmental model describing disease progression as illustrated in Figure 1a, where compartments describe population *densities* and are thus on the interval $[0, 1]$. We assume that the total population size (and density [2]) are fixed on the timescales of relevance.

Testing

Removal from I to R, including recovery, hospitalisation (perhaps requiring critical care), and death, occurs at a constant base rate γ_I. This rate is increased by *positive* testing of symptomatic individuals, expressing the fact that individuals who test positive will self-isolate or be placed in isolation earlier if tested earlier. We denote the positive testing rate τ, which is the fraction of the actual testing rate that produces positive results. The testing rate τ, given in (11h), is a function of two limiting parameters: the testing rate, τ_0, and the testing volume capacity, τ_∞ as introduced in [12].

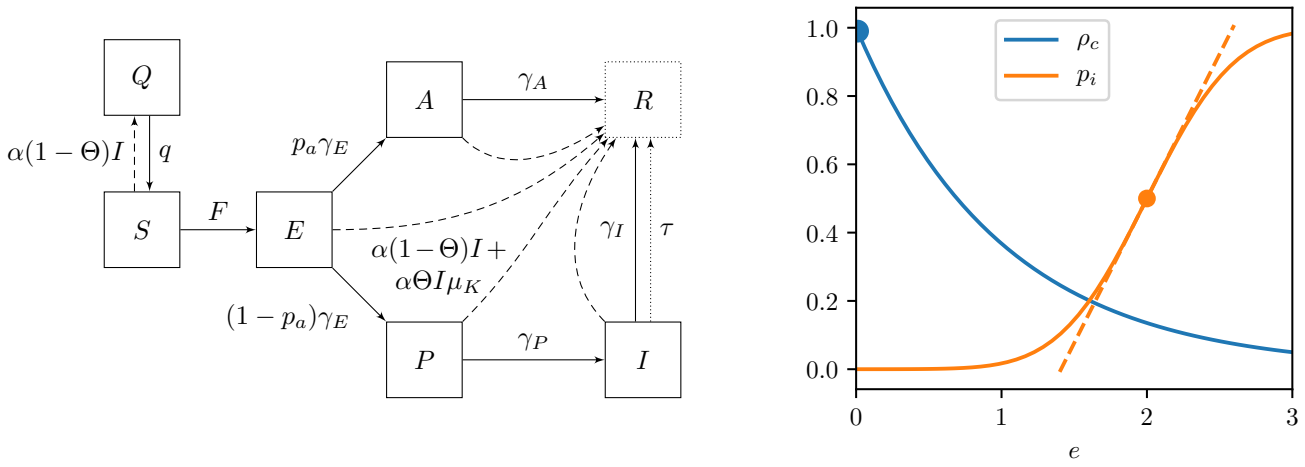
To understand these two parameters, consider two extreme cases. First, where only a handful of people need tests in a particular day, then the number of tests is not a limiting factor, but there will be some characteristic time that it takes to perform and process the tests, this is $1/\tau_0$. In the other extreme, if the entire population requires a test on a given day, there is a finite quantity of testing equipment and trained personnel that can administer the tests, therefore some maximum number density of people will be tested per day, τ_∞. These two limitations are particularly relevant in global pandemics, where testing capacity may be severely limited [9]. The parameters τ_0 and τ_∞ may be time-varying.

Contact tracing

For the sake of clarity, we assume throughout this section that no social intervention is imposed, thus $u_s(t) = 0$ for all t, and reinstate this factor only in the final model.

We introduce the notion of *exposure*, which quantifies the contact between interacting individuals. Naturally, exposure will depend on a multitude of factors, including the physical distance between the two individuals, the duration and nature of their contact, as well as the ambient conditions. The starting point for our model is the presumption that the developers of the digital tracing technology incorporate any combination of these factors into the application, and we are in possession of the average distribution of exposure to others in the population.

With this prelude, we denote by $c(e)$ the density of contacts at exposure level $e \in [0, \infty]$. We define zero exposure $e = 0$ to be the minimum level of exposure that the technology will detect, neglecting lower levels of exposure. For the purposes of parameterisation, we decompose $c(e)$ into a probability density $\rho_c(e)$ and a constant magnitude C, so $c(e) = C\rho_c(e)$ and $\int_0^\infty \rho_c(e)\,\mathrm{d}e = 1$. The parameter C is the number of contacts per unit time, encountered during a period of infectious duration. The density ρ_c describes how these contacts are distributed over exposure levels.



(a) (b)

Figure 1: ***(a)*** *Disease dynamics, testing and contact tracing framework, with rates and parameters defined in (11) and Tables 1 and 2. The dotted line illustrates testing, while the dashed lines illustrate contact tracing.* ***(b)*** *Examples of the contact exposure distribution $\rho_c(e)$ and the associated probability of infection $p_i(e)$. In our simulations we use the forms (12), where ρ_c is parametrised by a single parameter, $\rho_c(0) = 1/c_0$ (blue disk), and p_i is parametrised by two parameters, $p_i(e_{1/2}) = 1/2$ (orange disk) and $p_i'(e_{1/2}) = 1/(e_0\sqrt{\pi})$ (orange dashed line). For this illustration we set $c_0 = 1$, $e_{1/2} = 2$, and $e_0 = 2/3$.*

We associate each exposure level e with a corresponding probability of infection $p_i(e)$, representing the proportion of people who, at a given level of exposure throughout the infectious period (with respect to a nominal level of infectiousness, as we will describe), contract the disease. Note that $p_i(e)$ is a probability mass, not a probability density.

We highlight that, in general, we need not assume any specific forms for these two densities. Moreover, these distributions need not be arbitrary modeling choices, instead, they are to be collected (for example, the contact distribution C and ρ_c may be obtained directly from the contact tracing application) or deduced from separate studies (for example, the infection distribution p_i may be learned from application data in combination with virological studies). However, it seems sensible that ρ_c is monotonically decreasing, while p_i is monotonically increasing, examples of which we illustrate in Figure 1b. We henceforth make the latter assumption, which simplifies the presentation but may, in principle, be relaxed. Equipped with these preliminary quantities, we proceed to construct the contact tracing model.

First, we denote the proportion of the population that have adopted the contact tracing technology by $u_a(t)$. Upon a positive diagnosis, we have (for full adoption $u_a = 1$) the average distribution of $C\rho_c(e)$ contacts available to alert (per unit time per unit exposure). We define the notification threshold control $u_n(t) \in [0, \infty]$ such that only contacts with exposures above this control are informed of their exposure to the positively diagnosed individual. The proportion of contacts we alert, $f_c(u_n)$, is then given by

$$f_c = \int_{u_n(t)}^{\infty} \rho_c(e)\,\mathrm{d}e, \tag{1}$$

and the number of contacts we alert (per unit time) is Cf_c. Next, again retaining $u_a = 1$, the proportion of alerted contacts who would have been infected by a single newly diagnosed individual if all contacts were susceptible is denoted $\eta f_i(u_n)$, where

$$f_i = \frac{1}{f_c(u_n(t))} \int_{u_n(t)}^{\infty} \rho_c(e)p_i(e)\,\mathrm{d}e, \tag{2}$$

and η is a linear scale of infectiousness that will be determined from a self-consistency condition. It may be thought of by considering p_i as a probability with respect to a nominal level of infectiousness, and the scale η fixes this probability with respect to the instantaneous infectiousness. We then model the proportion of alerted contacts that were actually infected via $S\eta f_i$.

The number of contacts infected is more accurately described as an integral over the duration of the infectious period, however, the integral term is well approximated [3, 8] by the instantaneous integrand value (multiplied by the duration of infection), under the following assumption:

Assumption 1 (Timescale separation)**.** *The characteristic timescale of the duration of infectiousness is significantly smaller than the timescale of the population-wide disease dynamics.*

Assumption 1 is valid for many diseases, in particular COVID-19, as the typical infectiousness period is on the order of one week, while the timescale of population dynamics is several months or years (see Table 1). The multiple time-scales described by Assumption 1 justify treating the individual disease progression independently of the population disease dynamics (see for example [7, Ch. 7]).

We have introduced quantities describing contacts and their resulting infections that are already encapsulated in the disease dynamics. Thus, we must enforce self-consistency by ensuring that these quantities coincide. To this end, we define the instantaneous average transmission rate of symptomatic individuals assuming they were not removed via contact tracing but after testing, $\bar{\beta}$, by

$$\bar{\beta}(\tau) = \frac{\beta_P \gamma_P^{-1} + \beta_I (\gamma_I + \tau)^{-1}}{\gamma_P^{-1} + (\gamma_I + \tau)^{-1}}, \tag{3}$$

where we highlight the dependence on the time-varying τ. While the average transmission rate for a single individual ought to be the instantaneous transmission rate integrated over time, invoking Assumption 1, we approximate this by the instantaneous average (3).

The average transmission rate $\bar{\beta}$ represents the average number of infectious contacts per unit time (of an individual who eventually tests positive). This is the case with frequency-dependent transmission, however, this interpretation is without loss of generality, as we could adopt a density-dependent interpretation with the associated change in units [2]. This quantity is also represented by $C\eta f_i(0)$, thus self-consistency amounts to the constraint

$$C\eta f_i(0) = \bar{\beta}, \tag{4}$$

which determines the infectiousness scale η. In principle, C and ρ_c are to be obtained from the contact tracing application, while we derive $\bar{\beta}$ from clinical studies. Enforcing the constraint (4) then amounts to scaling the probability p_i by the factor η. For the probabilities to remain valid we require $\eta p_i(e) \leq 1$ for all e. Since we expect $p_i(e) \to 1$ as $e \to \infty$, and $\bar{\beta}$ varies monotonically with τ which we denote by $\bar{\beta}(\tau)$, we require the static compatibility condition

$$C f_i(0) \geq \max\left(\bar{\beta}(\tau_0), \bar{\beta}(\min(\tau_0, \tau_\infty))\right). \tag{5}$$

Intuitively, the compatibility condition (5) requires that the infectious contact profile introduced is capable of matching the infectious contact described by the disease dynamics.

The equality in (4) implicitly assumes that the contact tracing can detect all disease-transmitting contact. However, some disease transmission might be beyond the reach of digital contact tracing (e.g. indirect transmission via the environment [5]). This assumption can be relaxed, by replacing the right-hand side of (4) with the fraction of $\bar{\beta}$ that is detectable, which relaxes the right-hand side of the compatibility condition (5) by the same factor. For simplicity, we do not introduce this additional variable, and assume that other modes of transmission are negligible. We conclude by mentioning that the self-consistency condition (4) could, in fact, be used to help pin down the infection probability $p_i(e)$.

We now combine these elements to form the contact tracing component of the model. The contact tracing removal rate α is given by: the fraction of positive diagnoses per unit time using the digital application $u_a \tau(I) I$, the fraction of contacts encountered per unit time using the digital application $u_a(1-u_s)C$, the average infection time $[\gamma_P^{-1} + (\gamma_I + \tau)^{-1}]$, and the proportion of those contacts who are alerted $f_c(u_n)$.

The contact tracing removal rate is partitioned among *true positives*, those contacts alerted who were infected by the tested individual, and *false positives* who were instructed to isolate despite not being infected, in proportion with the tracing precision $\Theta := S\eta f_i$ and $(1-\Theta)$, respectively. We assume a well-mixed population (in accordance with the infection kinetics), such that the false positives are distributed uniformly within the population. The true positives are distributed among the infected and removed compartments, described by the proportions μ_K for $K \in \{E, A, P, I, R\}$. These proportions are derived in Section S1 B using the disease progression dynamics.

S1 B Deriving the μ_K probabilities

Our goal in this section is to derive the probabilities μ_K that a (true positive) traced contact is in compartment $K \in \{E, A, P, I, R\}$ at the time of tracing. To this end, we introduce the transition probability

$$P_{B,s_2|A,s_1} = \mathbb{P}(\text{individual in } B \text{ at } t = s_2 \mid \text{individual in } A \text{ at } t = s_1), \tag{6}$$

assuming the individual progresses along a continuous-time Markov chain described by the disease progression. Since the transition probabilities describe transitions of an individual throughout their disease progression, we leverage the time-scale separation assumption Assumption 1, from which we deduce time-homogeneity, namely

$$P_{B,s_2|A,s_1} = P_{B,s_2-s_1|A,0} =: P_{B|A}(s_2 - s_1). \tag{7}$$

Defining the time $t = 0$ as the time of the positive diagnosis of the tested case, our aim is to calculate the probability that a *contact* infected by this individual is in a compartment $K \in \{E, A, P, I, R\}$ at $t = t_0 \geq 0$, which we denote μ_K. The case of $t_0 = 0$ models immediate tracing, while $t_0 > 0$ captures delayed contact tracing. Denoting by $\mathbb{Q}$ the probability density of an individual infecting a contact (given that the individual tests positive at time $t = 0$), we find, using Bayes' law, and time-homogeneity, that

$$\begin{aligned}
\mu_K &= \mathbb{P}(\text{infected contact in } K \text{ at } t = 0) && \text{(8a)}\\
&= \int_0^\infty P_{K,t_0|E,-s}\, \mathbb{Q}(\text{infected contact at } t = -s)\, \mathrm{d}s && \text{(8b)}\\
&\propto \int_0^\infty P_{K|E}(t_0 + s) \sum_{J\in\{P,I\}} \beta_J P_{J,-s|I,0}\, \mathrm{d}s && \text{(8c)}\\
&= \int_0^\infty P_{K|E}(t_0 + s) \sum_{J\in\{P,I\}} \beta_J P_{I,0|J,-s} \frac{\mathbb{P}(\text{individual in } J \text{ at } t = -s)}{\mathbb{P}(\text{individual in } I \text{ at } t = 0)}\, \mathrm{d}s && \text{(8d)}\\
&\approx \int_0^\infty P_{K|E}(t_0 + s) \sum_{J\in\{P,I\}} \beta_J P_{I|J}(s) \frac{J}{I}\, \mathrm{d}s && \text{(8e)}\\
&=: \nu_K. && \text{(8f)}
\end{aligned}$$

where we have again used the time-scale separation from Assumption 1 to approximate the absolute probabilities. One might be concerned that, in calculating the integral over an infinite time horizon, the time-scale separation assumption breaks down. However, the long-time contributions are exponentially small due to this time-scale separation, and hence the accuracy of the approximation is retained.

Given the functions $P_{B|A}$, we may calculate the ν_K in (8f), from which the probabilities μ_K may be determined by normalising. The calculation of functions $P_{B|A}$ is straightforward, but, for the sake of completeness, is included in Appendix A.

The proportion ν_K defined in (8e) may be written as

$$\nu_K = \sum_{J\in\{P,I\}} \beta_J \frac{J}{I}\, p_{K,J}, \qquad\qquad p_{K,J} = \int_0^\infty P_{K|E}(t_0 + s) P_{I|J}(s)\, \mathrm{d}s, \tag{9}$$

where explicit expressions are available for $p_{K,J}$ but are too unwieldy to reproduce here. The quantities $p_{K,J}$ depend on the instantaneous transition rates, and therefore on the state I through false positives $\alpha(1 - \Theta)I$ and the testing $\tau(I)$ for the I compartment. The probabilities μ_K, obtained by normalising ν_K, take the form

$$\mu_K = \frac{\nu_K}{\sum_{K\in\{E,A,P,I,R\}} \nu_K} = \frac{\beta_I p_{K,I} + \beta_P p_{K,P} P/I}{\beta_I \left(\sum_K p_{K,I}\right) + \beta_P \left(\sum_K p_{K,P}\right) P/I}. \tag{10}$$

The probabilities μ_K depend on the state via the ratio P/I, as well as through the dependence of $p_{K,J}$ on I, and we denote this via $\mu_K(I, P/I)$.

S1 C Model equations and default parameter values

The contact tracing component of our compartmental model (and the associated analysis) underpins the key contribution in this work. We briefly summarise our construction before presenting the full model equations.

We presume that we are in possession of the real-world distribution of contacts at different exposure levels, and assume an associated probability of infection at each exposure level. We introduce the control $u_n \in [0, \infty]$, which is the notification threshold for instructing an individual to isolate: only contacts who were exposed to the positively tested individual at or above the level u_n are directed to isolate. Moreover, we introduce the control u_a, representing the proportion of the population that has adopted the digital contact tracing application. Both controls u_n and u_a are allowed to vary with time. Based on these quantities, we calculate the tracing rate α, and the tracing precision Θ.

Upon reintroducing the social intervention measures via u_s, the model outlined above may be written as an initial-value problem, comprising the system of ODEs and initial conditions given by

$$\dot{S} = -FS + qQ \qquad - \alpha(1-\Theta)IS, \qquad S(0) = S_0, \tag{11a}$$
$$\dot{Q} = -qQ \qquad + \alpha(1-\Theta)IS, \qquad Q(0) = Q_0, \tag{11b}$$
$$\dot{E} = FS - \gamma_E E \qquad - \alpha(1-\Theta)IE - \alpha\Theta I\mu_E, \qquad E(0) = E_0, \tag{11c}$$
$$\dot{A} = p_a\gamma_E E - \gamma_A A \qquad - \alpha(1-\Theta)IA - \alpha\Theta I\mu_A, \qquad A(0) = A_0, \tag{11d}$$
$$\dot{P} = (1-p_a)\gamma_E E - \gamma_P P - \alpha(1-\Theta)IP - \alpha\Theta I\mu_P, \qquad P(0) = P_0, \tag{11e}$$
$$\dot{I} = \gamma_P P - (\gamma_I + \tau)I \qquad - \alpha(1-\Theta)I^2 - \alpha\Theta I\mu_I, \qquad I(0) = I_0, \tag{11f}$$

where

$$F(u_s, A, P, I) = [1-u_s]\left(\beta_A A + \beta_P P + \beta_I I\right), \tag{11g}$$
$$\tau(I) = \min(\tau_0, \tau_\infty / I), \tag{11h}$$
$$\alpha(u_s, u_a, u_n, I) = u_a^2 \tau [1-u_s] C f_c(u_n)[\gamma_P^{-1} + (\gamma_I + \tau)^{-1}], \tag{11i}$$
$$\Theta(u_n, S, I) = \frac{S f_i(u_n)}{C f_i(0)} \frac{\beta_P \gamma_P^{-1} + \beta_I(\gamma_I + \tau)^{-1}}{\gamma_P^{-1} + (\gamma_I + \tau)^{-1}}, \tag{11j}$$

where we have not written down the R dynamics, as these are recoverable via conservation. Should one wish to account for the number of hospitalised, dead, etc. additional compartments to accumulate these quantities may be added to the topology. For the sake of simplicity, in this work we focus on the central dynamics of the model, and leave the study of other quantities to future work.

We consider only the scenario in which traced contacts are isolated but not tested for the disease, and we exclude tracing or isolation of secondary contacts (contacts of contacts). Extensions to include such strategies are possible within the current framework, but omitted in this work for the sake of simplicity.

As mentioned, the distributions $\rho_c(e)$ and $p_i(e)$ may be learned from application data. For the sake of concreteness, in this study we specify the parameterised forms

$$\rho_c(e) = \frac{1}{c_0} \exp\left(-\frac{e}{c_0}\right), \qquad p_i(e) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{e - e_{1/2}}{e_0}\right) + 1\right], \tag{12}$$

where erf is the error function $\operatorname{erf}(x) = \int_0^x (2/\sqrt{\pi}) \exp(-y^2)\, \mathrm{d}y$, and we have introduced three parameters: the contact variability c_0, the typical infectious exposure level $e_{1/2}$, and infectious uncertainty e_0. These forms are illustrated in Figure 1b, and give explicit forms for the integrals (1) and (2), namely

$$f_c = \exp(-u_n(t)/c_0), \tag{13}$$
$$f_i = \frac{1}{2f_c}\left\{ \mathrm{e}^{[e_0/(2c_0)]^2 - e_{1/2}/c_0}\left[1 - \operatorname{erf}\left(\frac{e_0}{2c_0} + \frac{u_n(t) - e_{1/2}}{e_0}\right)\right] + \mathrm{e}^{-u_n(t)/c_0}\left[1 - \operatorname{erf}\left(\frac{e_{1/2} - u_n(t)}{e_0}\right)\right]\right\}. \tag{14}$$

Unless stated otherwise, numerical simulations in this work employ the default parameter values listed in Table 1 and the control parameters listed in Table 2. Numerical simulations serve primarily to demonstrate the presented contact tracing framework. Therefore, specific parameter values are not of core importance for the results presented in this paper. Nevertheless, it is important to choose parameter values that capture key characteristics of the modeled disease for the numerical simulations to be meaningful. In the following, we detail how the default parametrisation has been chosen.

The proportion of asymptomatic cases has been chosen as 20% in line with [10]. The rates γ_K, $K \in \{E, A, P, I\}$ determine typical times spend at different stages of infectiousness. $\gamma_E = 1/3.5$ day^{-1} describes an average of 3.5 days from infection to the beginning of the infectious period. Taking an average incubation time (the time between infection and the onset of symptoms) of 5.5 days [5, 10], implies that individuals are infectious on average 2 days before becoming symptomatic, as characterised by $\gamma_P = 1/2$ day^{-1}. γ_A and γ_I determine how long individuals remain infectious. These rates are difficult to determine since infectiousness durations are difficult to measure. The parameter γ_I has the added complexity that it combines the end of the infectious period with typical self-isolation or hospitalisation. To overcome this difficulty, we note that when the typical durations of infectious periods are changed by varying γ_K, $K \in \{A, P, I\}$ but the transmission rates β_K, $K \in \{A, P, I\}$ are kept constant, the basic reproduction number $\mathcal{R}_0$ (number of secondary infections) is modified. This is undesirable since the $\mathcal{R}_0$ is easier to quantify than the γ_K parameters. Motivated by reference [5], we determine the

Parameter	Description	Default value	Unit
β_A	transmission rate with asymptomatic	$0.2\gamma_A/p_a$	day^{-1}
β_P	transmission rate with presymptomatic	$1.3\gamma_P/(1-p_a)$	day^{-1}
β_I	transmission rate with symptomatic	$1.2\gamma_I/(1-p_a)$	day^{-1}
γ_E	incubation rate	$1/3.5$	day^{-1}
γ_A	asymptomatic recovery rate	$1/5$	day^{-1}
γ_P	symptom development rate	$1/2$	day^{-1}
γ_I	testable removal rate	$1/4$	day^{-1}
τ_0	testing rate (for a single individual)	2	day^{-1}
τ_∞	maximum testing capacity per unit time	∞	day^{-1}
q	quarantine removal rate	$1/14$	day^{-1}
p_a	proportion of asymptomatic infections	0.2	—
C	average number of contacts per unit time	20	day^{-1}
c_0	contact variability	1	exposure
$e_{1/2}$	typical infectious exposure level	3	exposure
e_0	infectious exposure variation	2/3	exposure
t_0	contact tracing delay	0	day
Σ_0	initial infectious proportion	10^{-6}	—
E_0	initial latent exposed proportion	$\Sigma_0 c_E/(c_E+c_A+c_P+c_I)$	—
A_0	initial asymptomatic proportion	$\Sigma_0 c_A/(c_E+c_A+c_P+c_I)$	—
P_0	initial presymptomatic proportion	$\Sigma_0 c_P/(c_E+c_A+c_P+c_I)$	—
I_0	initial symptomatic proportion	$\Sigma_0 c_I/(c_E+c_A+c_P+c_I)$	—
Q_0	initial quarantined proportion	$\Sigma_0 c_Q/(c_E+c_A+c_P+c_I)$	—
S_0	initial susceptible proportion	$S^* - \Sigma_0(1+c_Q)$	—
S^*	disease-free susceptible proportion	1	—

Table 1: *Default parameter values for simulations. Values c_K for $K \in \{E, A, P, I, Q\}$ are defined in Section S2 A.*

Control	Description	Range	Value	Unit
u_s	social intervention measures	$[0, 1]$	0	—
u_a	contact tracing adoption fraction	$[0, 1]$	1	—
u_n	notification threshold	$[0, \infty]$	0	exposure

Table 2: *Control parameters.*

transmission rates β_K, $K \in \{A, P, I\}$ as functions of γ_K, $K \in \{A, P, I\}$ together with given contributions of each compartment to $\mathcal{R}_0$. In particular, we assumed that asymptomatic cases contribute 0.2, presymptomatic cases 1.3, and symptomatic cases 1.2 secondary infections to the total $\mathcal{R}_0 = 2.7$ (see [1]). This approach ensures that varying rates γ_K does not change the reproduction number.

The testing rate $\tau_0 = 2\ \text{day}^{-1}$ corresponds to it taking on average half a day from the onset of symptoms until the individual is tested, isolated, and contact traced. The choice of this value has been motivated by plans to incorporate the possibility for individuals to automatically request a test through the digital contact tracing application. For the early phase of the COVID-19 pandemic when the testing infrastructure still had to be established in most countries, this short duration is likely to be overly optimistic. We note, however, that it is more pessimistic than the analysis in [5] where it was assumed that case isolation happens immediately after symptom onset. An implicit feature of our formulation of testing in the model is that, even for non-saturated testing capacities, a fraction $(\gamma_I/(\tau + \gamma_I))$ of symptomatic cases, progresses to the removed compartment without being tested and contact traced. These cases correspond to symptomatic individuals that are never identified (for instance, those that recover quickly and do not request a test), and cause an implicit case isolation and contact tracing inefficiency.

In our simulations, the epidemic commences with a small proportion of infected individuals $\Sigma_0 = 10^{-6}$. The precise value is of little consequence for the presented results. The contact tracing delay t_0 is kept at zero throughout the main paper while non-zero delays are explored in Section S3 D. Other parameters are explored throughout the paper (τ_∞, u_s, u_a, u_n) or do not matter much for the presented results. The parameters q and C primarily scale the number of individuals in the Q compartment. We note that the number of contacts per day, C, affects the contact tracing precision Θ. This must necessarily be the case because the number of secondary infections is determined by the disease dynamics. Therefore, when the number

of contacts per day is varied by changing C, the infection probabilities per contact are scaled appropriately to preserve the same number of secondary infections (see self-consistency constraint (5)). As a result, changing C in the model preserves the number of infectious contacts and only changes the number of contacts that do not lead to an infection. This manifests itself in the model through the fact that C cancels out in the true positive contact tracing removal rate ($\alpha\Theta$) but is retained in the false positive removal rate, since the number of false positive notified contacts increases with the number of not infectious contacts.

S2 Mathematical analysis

S2 A Evaluation at the disease-free equilibrium

In this section, we study the system of ODEs (11). As we will describe, the system is singular at the disease-free equilibrium (DFE), and the aim of this section is to "regularise" this singularity by providing a meaningful way to evaluate the system (and its Jacobian) at the DFE.

We consider solutions of (11) for initial conditions in $\Omega \subset \mathbb{R}_+^6$ defined by

$$\Omega = \{(S, Q, E, A, P, I) \in \mathbb{R}_+^6 \mid S + Q + E + A + P + I \leq 1\}, \tag{15}$$

where $\mathbb{R}_+$ denotes the non-negative reals. For convenience, we denote the system (11) by $\dot{x} = f(x, t)$ where f is the right-hand side of (11). The function f has a removable singularity for all $x \in \mathcal{D}$ given by

$$\mathcal{D} = \{E = A = P = I = 0\} \cap \Omega, \tag{16}$$

stemming from the probabilities $\mu_K(I, P/I)$. The singularity may be removed by defining $f(x, t) = (qQ, -qQ, 0, \ldots, 0)$ for all $x \in \mathcal{D}$ and all t, as the singular terms appear only in the form $I\mu_K$ which vanishes as $x \to \mathcal{D}$ since $0 \leq \mu_K \leq 1$.

Not all trajectories with initial conditions in Ω remain within Ω. For example, if $E_0 = A_0 = 0$ while $I_0 > 0$, then it follows from (11d) that $\dot{A}(0) < 0$ and hence $A < 0$ at least for some small initial transient. These unrealistic negative densities could be avoided, by, for example, replacing μ_K with $\mu_K g(K)$ for a sigmoidal g such that $g(0) = 0$ and $g(\epsilon) \approx 1$ for $\epsilon \ll 1$. However, as we will demonstrate, this modification is not necessary for the class of initial conditions we are interested in: when the infectious densities respect the disease dynamics in some sense, the trajectories remain in Ω. It therefore suffices, for our purposes, to analyse f in Ω.

The right-hand side f has a uniformly bounded gradient (with respect to the state) on $\Omega \setminus \mathcal{D}$, however, on $\mathcal{D}$ the singularity persists in the gradient and is not removable, as we will discuss. Nonetheless, f is Lipschitz in Ω with respect to the state, and inherits its time regularity from the controls u_s, u_a and u_n, which we assume are time-continuous for the sake of simplicity, however, this assumption may be relaxed. Similarly, f is continuously differentiable with respect to the parameters. Classical results [6, ch. 1] then guarantee that a unique solution of the system of ODEs exists and exhibits continuous dependence on initial conditions and parameters.

These conditions are not sufficient for our purposes, since we want to analyse the local behaviour at the disease-free equilibrium (DFE), which is in $\mathcal{D}$, requiring the existence of the Jacobian of f. It is instructive to demonstrate the nature of this indefiniteness as this informs our approach to overcome the singularity. The troublesome contributions to the Jacobian are the terms $I\mu_K(I, P/I)$ whose derivatives depend on state exclusively through the ratio P/I. We may approach $\mathcal{D}$ along sets $\ell(E) = \{P = c_P E, I = c_I E\}$ as $E \to 0$, such that $P/I = c_P/c_I$ takes any arbitrary value, thus this term is irremovably singular on $\mathcal{D}$. That f is Lipschitz can be seen by noting that, despite not having a unique value, the Jacobian of f is uniformly bounded since all the coefficients of P and I in the expression (10) for μ_K are strictly positive.

Our approach is to evaluate this ratio P/I by demonstrating that, in the vicinity of $\mathcal{D}$, the ratio of any two states is asymptotically constant. This observation guides the evaluation of the Jacobian in $\mathcal{D}$: terms of the form P/I can be replaced by constants that are determined from the dynamics.

We begin by considering the dynamics in the vicinity of $\mathcal{D}$, when all compartments except S contain a density of order $\mathcal{O}(\epsilon)$, while $S = S^* - \mathcal{O}(\epsilon) = \mathcal{O}(1)$ for some asymptotically small ϵ, and thus this pertains to trajectories in the vicinity of $\mathcal{D}$. We seek a solution in the asymptotic limit as $\epsilon \to 0$, and scale the system state via

$$S = S^* - \epsilon\hat{S}, \qquad Q = \epsilon\hat{Q}, \qquad E = \epsilon\hat{E}, \qquad A = \epsilon\hat{A}, \qquad P = \epsilon\hat{P}, \qquad I = \epsilon\hat{I}. \tag{17}$$

We now expand each quantity asymptotically in powers of ϵ, whereby, dropping the hats to simplify notation, the leading-order

system takes the form

$$\dot{S} = [1 - u_s(t)]\,(\beta_A A + \beta_P P + \beta_I I)\,S^* - qQ + \alpha(1-\Theta)S^* I, \tag{18a}$$
$$\dot{Q} = -qQ + \alpha(1-\Theta)S^* I, \tag{18b}$$
$$\dot{E} = [1 - u_s(t)]\,(\beta_A A + \beta_P P + \beta_I I)\,S^* - \gamma_E E - \alpha\Theta I \mu_E(0, P/I), \tag{18c}$$
$$\dot{A} = p_a \gamma_E E - \gamma_A A - \alpha\Theta I \mu_A(0, P/I), \tag{18d}$$
$$\dot{P} = (1-p_a)\gamma_E E - \gamma_P P - \alpha\Theta I \mu_P(0, P/I), \tag{18e}$$
$$\dot{I} = \gamma_P P - (\gamma_I + \tau(0))I - \alpha\Theta I \mu_I(0, P/I), \tag{18f}$$

and we write $\tau = \tau(0)$ for

$$\tau(0) = \tau_0 \operatorname{sign}(\tau_\infty) = \begin{cases} \tau_0, & \tau_\infty > 0, \\ 0, & \tau_\infty = 0, \end{cases} \tag{19}$$

and thus α and Θ do not depend on the state, and are given by

$$\alpha = u_a^2 \tau(0)[1-u_s] C f_c(u_n)[\gamma_P^{-1} + (\gamma_I + \tau(0))^{-1}], \tag{20}$$
$$\Theta = \frac{S^* f_i(u_n)}{C f_i(0)} \frac{\beta_P \gamma_P^{-1} + \beta_I(\gamma_I + \tau(0))^{-1}}{\gamma_P^{-1} + (\gamma_I + \tau(0))^{-1}}. \tag{21}$$

Assuming quasi-steady control inputs u_s, u_a and u_n, we seek an exponential solution to the leading-order system (18) of the form

$$S = c_S \mathrm{e}^{\lambda t}, \qquad Q = c_Q \mathrm{e}^{\lambda t}, \qquad E = c_E \mathrm{e}^{\lambda t}, \qquad A = c_A \mathrm{e}^{\lambda t}, \qquad P = c_P \mathrm{e}^{\lambda t}, \qquad I = c_I \mathrm{e}^{\lambda t}. \tag{22}$$

Substituting (22) into (18) we obtain the system of algebraic equations

$$\lambda c_S = [1 - u_s]\,(\beta_A c_A + \beta_P c_P + \beta_I c_I)\,S^* - q c_Q + \alpha(1-\Theta)S^* c_I, \tag{23a}$$
$$\lambda c_Q = -q c_Q + \alpha(1-\Theta)S^* c_I, \tag{23b}$$
$$\lambda c_E = [1 - u_s]\,(\beta_A c_A + \beta_P c_P + \beta_I c_I)\,S^* - \gamma_E c_E - \alpha\Theta c_I \mu_E(0, c_P/c_I), \tag{23c}$$
$$\lambda c_A = p_a \gamma_E c_E - \gamma_A c_A - \alpha\Theta c_I \mu_A(0, c_P/c_I), \tag{23d}$$
$$\lambda c_P = (1-p_a)\gamma_E c_E - \gamma_P c_P - \alpha\Theta c_I \mu_P(0, c_P/c_I), \tag{23e}$$
$$\lambda c_I = \gamma_P c_P - (\gamma_I + \tau(0))c_I - \alpha\Theta c_I \mu_I(0, c_P/c_I). \tag{23f}$$

System (23) is homogeneous with respect to $(c_S, c_Q, c_E, c_A, c_P, c_I)$, and therefore we take $c_I = 1$ without loss of generality, and thus we have seven equations for seven unknowns. Equations (23c–23f) decouple from (23a) and (23b), and may be solved independently, and the solution plugged into (23b) followed by (23a) to determine c_Q and then c_S.

For some parameters, it turns out that $\lambda \leq -q$. In this regime, the disease is eradicated faster than the quarantined population in Q returns to compartment S, and thus the assumption that Q and the infectious compartments are of the same order breaks down. The leading-order dynamics in this regime describe a new time-scale separation, whereby Q acquires an impulse of susceptibles which return exponentially. For the sake of simplicity, and since the region in parameter space where this phenomenon is observed is not significant (requiring nearly full adoption rate u_a and high social intervention u_s), we do not discuss this regime further.

Solving (23) provides a wealth of structural information about the system. The sign of λ predicts whether the system will grow or decay exponentially in time. Crucially, all quantities exhibit the same exponential evolution, and thus the ratio of any two is constant. This motivates our replacement of the ratio P/I by $c_P/c_I = c_P$ when evaluating the Jacobian in $\mathcal{D}$.

We conclude by discussing how this analysis influences the choice of initial conditions. As mentioned at the beginning of this section, solutions depend continuously on initial conditions and parameters. From this observation we might expect similar results for sufficiently similar initial conditions. However, in the initial outbreak dynamics, the number of infected individuals may grow exponentially, therefore small initial discrepancies will be amplified and a realistic choice of initial conditions is essential. Moreover, unrealistic dynamics are observed during a short initial transient for general initial conditions. For example, it is tempting, for the sake of simplicity, to impose uniform initial infected conditions $E_0 = A_0 = P_0 = I_0$. However, as illustrated in Figure 2, this results in an initial dip in the total number of infected individuals $\Sigma = E + A + P + I$ despite the fact that a strong outbreak occurs. The source of this unrealistic dynamic is that the uniform distribution is

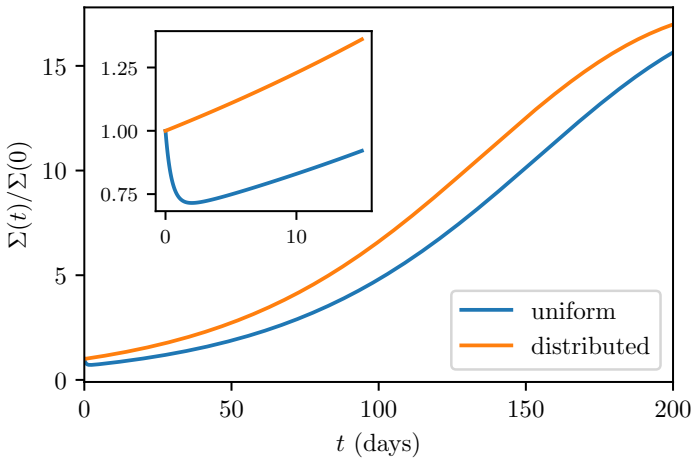


Figure 2: *The number of infected individuals* $\Sigma = E + A + P + I$ *during an initial transient for* $\Sigma(0) = 10^{-4}$ *and* $u_a = 0.7$, *where the infected and quarantine compartments are initiated uniformly* ($E_0 = A_0 = P_0 = I_0 = Q_0$) *or distributed according to* (23).

not obtained via evolution of the disease dynamics, and so some transient is needed for the disease to adopt the distribution dictated by the dynamics. While strictly speaking any non-negative initial condition may arise via population migration, the case of greatest interest in this study is those initial conditions that respect the dynamics.

To remedy this, the initial distribution is to be governed by the solution of (23). In Figure 2 we see that the same infection proportion $\Sigma(0)$ distributed among the infected compartments according to (23) eliminates the unrealistic initial dip. This observation has the additional advantage that the choice of initial conditions is reduced from six degrees of freedom to just one, and we choose to fix the proportion of infections in the population at the initial time $\Sigma(0) = E_0 + A_0 + P_0 + I_0$. This subset of initial conditions alleviates any concern about negative densities: the asymptotic analysis demonstrates that all densities remain positive.

To appropriately model the deployment of a digital contact tracing application, one needs to specify a time-varying adoption profile $u_a(t)$ with $u_a(0) = 0$. In this study, we consider only constant u_a profiles, neglecting the influence of time-varying adoption rates.

S2 B Approximations

In this section, we take the leading-order system derived in Section S2 A, relevant when the infection proportion is asymptotically small, and deduce asymptotic approximations of many of the quantities of interest. We have already seen how the asymptotic insight allows us to significantly reduce the degrees of freedom in choosing initial conditions, and allows us to evaluate the Jacobian of the system at the DFE. There are further benefits to be gained by calculating the quantities of interest: providing insight into the dominant forces driving the disease dynamics including closed-form expressions that elucidate dependence, validating the numerical implementation, and even replacing numerical simulations when these are excessively time-consuming, or beyond reach, as is the case in Figure 4 right of the main paper, where the time horizon is infinite.

We begin by approximating the proportion of the population that has been infected by time t, given by

$$S^* - S - Q \sim (c_S - c_Q)\mathrm{e}^{\lambda t}. \tag{24}$$

It follows that the time t to reach a proportion of infections p is given by

$$t \sim \frac{1}{\lambda} \log\left(\frac{p}{c_S - c_Q}\right). \tag{25}$$

We find the approximation (25) agrees well with the full numerical simulations in Figure 2 left. Importantly, the asymptotic results demonstrates that the range of τ_∞ presented is representative of the entire range of τ_∞ values.

One factor contributing to the cost of contact tracing is the quantity of susceptibles quarantined (occupying the Q compartment) during the period $t \in [0, T]$. This is proportional to the integral

$$\int_0^T Q \,\mathrm{d}t \sim c_Q \frac{\mathrm{e}^{\lambda T} - 1}{\lambda}. \tag{26}$$

When there is no outbreak $\lambda < 0$, and the integral (26) has a limit as $T \to \infty$, namely

$$\int_0^\infty Q \,\mathrm{d}t = -\frac{c_Q}{\lambda}. \tag{27}$$

We find that the approximation (27) agrees excellently with full numerical solutions in Figure 4 left. For this reason, and because we choose an infinite time horizon, the asymptotic approximation is used to produce the shaded region in Figure 4 right, as the numerical simulations are prohibitively time-consuming.

S2 C Final size relation

It is of great interest to know how many people will ultimately be infected without requiring full numerical simulations. In this section, we seek a relation that quantifies the final number of susceptible in the population after the epidemic has passed $S(\infty)$, in the absence of contact tracing $u_a \equiv 0$ and for constant intervention u_s. To obtain such a final size relation, we note that, since the model (11) admits only the disease-free equilibrium (DFE), the infectious states must vanish as $t \to \infty$. Integrating (11a) and substituting (11g), we find that

$$\log \frac{S(\infty)}{S_0} = -\int_0^\infty F \,\mathrm{d}t = -(1-u_s)\int_0^\infty \beta_A A(t) + \beta_P P(t) + \beta_I I(t) \,\mathrm{d}t. \tag{28}$$

We now use the disease dynamics (11) to evaluate the integral terms in (28), whereby

$$S(\infty) - S_0 - E_0 = \int_0^\infty \dot{S} + \dot{E} \,\mathrm{d}t = -\gamma_E \int_0^\infty E \,\mathrm{d}t, \tag{29a}$$

$$-A_0 = \int_0^\infty \dot{A} \,\mathrm{d}t \qquad = p_a \gamma_E \int_0^\infty E \,\mathrm{d}t - \gamma_A \int_0^\infty A \,\mathrm{d}t, \tag{29b}$$

$$-P_0 = \int_0^\infty \dot{P} \,\mathrm{d}t \qquad = (1-p_a)\gamma_E \int_0^\infty E \,\mathrm{d}t - \gamma_P \int_0^\infty P \,\mathrm{d}t, \tag{29c}$$

$$-I_0 = \int_0^\infty \dot{I} \,\mathrm{d}t \qquad = \gamma_P \int_0^\infty P \,\mathrm{d}t - \gamma_I \int_0^\infty I \,\mathrm{d}t - \int_0^\infty \tau(I) I \,\mathrm{d}t. \tag{29d}$$

The final term in (29d) admits the bound

$$0 \le \int_0^\infty \tau(I) I \,\mathrm{d}t \le \tau(0) \int_0^\infty I \,\mathrm{d}t. \tag{30}$$

By successive substitution in (29), we find that all the integral terms on the right-hand side of (28) may be expressed in relation to the final size term $S(\infty)$ and the problem constants (initial conditions and parameters), which provides the final size inequality

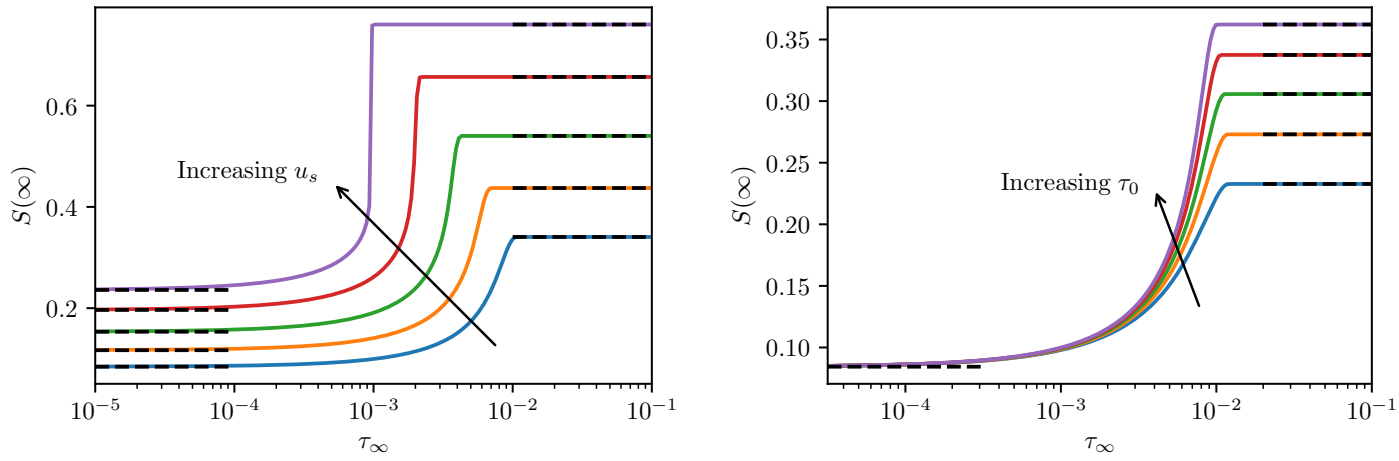
$$\begin{aligned} 0 \le &-\log \frac{S(\infty)}{S_0} - [S_0 + E_0 - S(\infty)]\mathcal{R}_0|_{u_a=0} - (1-u_s)\left(A_0 \frac{\beta_A}{\gamma_A} + P_0 \frac{\beta_P}{\gamma_P} + (P_0 + I_0)\frac{\beta_I}{\gamma_I + \tau(0)}\right) \\ &\le (1-u_s)\frac{\beta_I}{\gamma_I + \tau(0)}\left[(1-p_a)(S_0 + E_0 - S(\infty)) + P_0 + I_0\right]\frac{\tau(0)}{\gamma_I}. \end{aligned} \tag{31}$$

Since we expect the outbreak to start when $S_0 \approx 1$ and all other initial conditions are close to zero, we deduce the approximate relation

$$0 \lessapprox -\log S(\infty) - [1 - S(\infty)]\mathcal{R}_0|_{u_a=0} \lessapprox (1-u_s)\frac{\beta_I}{\gamma_I + \tau(0)}(1-p_a)(1-S(\infty))\frac{\tau(0)}{\gamma_I}. \tag{32}$$

With vanishingly slow testing $\tau_0 \ll 1$, we expect the contribution of the testing term to be negligible, whereby both bounds (32) converge to zero to provide the estimate

$$\log S(\infty) + [1 - S(\infty)]\mathcal{R}_0|_{\tau(0)=0} \approx 0. \tag{33}$$



(a) $u_s \in \{0, 0.1, 0.18, 0.25, 0.3\}$.

(b) $\tau_0 \in \{0.5, .8, 1.2, 1.9, 3\}$.

Figure 3: *The proportion of individual remaining uninfected after the epidemic $S(\infty)$ as a function of the testing capacity τ_∞ for various values of the (a) intervention impact u_s; (b) testing rate τ_0. The black dashed lines show the limiting approximations derived in Section S2 C.*

For a vanishingly small testing capacity $\tau_\infty \ll 1$, testing is quickly saturated in a significant outbreak, and the value of τ_0 has little influence. By this reasoning, similar dynamics may be obtained using a vanishingly small τ_0, and thus we expect the approximation (33) to be accurate.

Similarly, when the maximum testing capacity is never exceeded, $I(t) \leq \tau_\infty/\tau_0$ for all t, then the upper bound in (30) is obtained with equality, and following analogous calculations, we obtain the final-size estimate

$$\log S(\infty) + [1 - S(\infty)]\mathcal{R}_0|_{\tau(0)=\tau_0} \approx 0. \tag{34}$$

In Figure 3 we plot $S(\infty)$ as a function of τ_∞ for various social intervention factors u_s and testing rates τ_0. The black dashed lines show the limiting cases as $\tau_\infty \to 0$ and $\tau_\infty \to \infty$, given in (33) and (34). We find good agreement with the full numerical simulation, which, while justifying that the simulations were carried out to sufficiently large times, took significant computational effort.

S2 D Basic reproduction number

In this section, we aim to calculate the basic reproduction number of (11) using the next-generation method [4, 11], and the analysis of Section S2 A which allows us to evaluate the Jacobian at the DFE. To this end, we write the system state in vector form $\mathcal{Y} = (S, Q, E, A, P, I)$. The state vector may be decomposed into (S, Q), the uninfected components, and the remaining compartments, all describing infected states at different stages of illness progression. The regularised system dynamics (see Section S2 A) may then be decomposed via

$$\dot{\mathcal{Y}} = \mathcal{F}(\mathcal{Y}) - \mathcal{V}(\mathcal{Y}), \tag{35}$$

where the rate $\mathcal{F}$ describes new infections, while $\mathcal{V}$ incorporates all other transmissions between compartments (see [11] for a detailed explanation of the decomposition in a general setting). In particular,

$$\mathcal{F}(\mathcal{Y}) = \begin{pmatrix} 0 \\ 0 \\ FS \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathcal{V}(\mathcal{Y}) = \begin{pmatrix} FS - qQ \\ qQ \\ \delta E \\ -p_a\gamma_E E + \gamma_A A \\ -(1-p_a)\gamma_E E + \gamma_P P \\ -\gamma_P P + (\gamma_I + \tau)I \end{pmatrix} + \begin{pmatrix} \alpha(1-\Theta)IS \\ -\alpha(1-\Theta)IS \\ \alpha(1-\Theta)IE + \alpha\Theta I\mu_E(I, P/I) \\ \alpha(1-\Theta)IA + \alpha\Theta I\mu_A(I, P/I) \\ \alpha(1-\Theta)IP + \alpha\Theta I\mu_P(I, P/I) \\ \alpha(1-\Theta)I^2 + \alpha\Theta I\mu_I(I, P/I) \end{pmatrix}. \tag{36}$$

We highlight the unique (up to a multiplicative scale) disease-free equilibrium (DFE) $\mathcal{Y}^* = (S^*, 0, \ldots, 0)$, around which

we linearise the system (35), to obtain the Jacobian matrices evaluated at the DFE,

$$D\mathcal{F}(\mathcal{Y}^*) = \begin{pmatrix} 0 & 0 \\ 0 & F \end{pmatrix}, \qquad D\mathcal{V}(\mathcal{Y}^*) = \begin{pmatrix} M_1 & M_2 \\ 0 & V \end{pmatrix}, \tag{37}$$

where $F, V \in \mathbb{R}^{4\times4}$ matrices, $M_2 \in \mathbb{R}^{2\times4}$, and $M_1 \in \mathbb{R}^{2\times2}$ associated with the state decomposition above. In particular,

$$F = (1-u_s)S^* \begin{pmatrix} 0 & \beta_A & \beta_P & \beta_I \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} \gamma_E & 0 & C_{E,P} & C_{E,I} \\ -p_a\gamma_E & \gamma_A & C_{A,P} & C_{A,I} \\ -(1-p_a)\gamma_E & 0 & \gamma_P + C_{P,P} & C_{P,I} \\ 0 & 0 & -\gamma_P + C_{I,P} & \gamma_I + \tau(0) + C_{I,I} \end{pmatrix}, \tag{38}$$

where $\tau(0)$ is given by (19) and the contact tracing terms are given by the shorthand

$$C_{K,I} = \alpha\Theta[\mu_K(0, c_P) - c_P\mu_K'(0, c_P)], \qquad C_{K,P} = \alpha\Theta\mu_K'(0, c_P), \tag{39}$$

for $K \in \{E, A, P, I\}$, where primes denote differentiation with respect to the state ratio, c_P is deduced in Section S2 A, and $\alpha\Theta$ is given by

$$\alpha\Theta = S^* u_a^2 \tau(0)(1-u_s) f_c(u_n) \left[\beta_P\gamma_P^{-1} + \beta_I(\gamma_I + \tau(0))^{-1}\right] \frac{f_i(u_n)}{f_i(0)}. \tag{40}$$

The matrix F describes the rates at which infected individuals produce secondary infections, while the matrix V describes the transmission rates of infected individuals amongst other infected compartments *excluding secondary infections* but including infection removal, such that V^{-1} describes the average residence durations in each infected compartment [11]. The next-generation matrix FV^{-1} then describes the expected number of new infections produced by infected individuals in each infected compartment.

This leads naturally to defining the basic reproduction number $\mathcal{R}_0$ describing the average number of secondary infections generated by a primary infection in a susceptible population, which is given by the spectral radius of the next generation matrix FV^{-1},

$$\mathcal{R}_0 = \rho(FV^{-1}) = \max\{|\lambda| : \lambda \text{ eigenvalue of } FV^{-1}\}. \tag{41}$$

It is well-known [11] that $\mathcal{R}_0$ is a threshold parameter: if $\mathcal{R}_0 > 1$ the system is unstable, and the introduction of infected individuals leads to a disease outbreak, whereas if $\mathcal{R}_0 < 1$ the introduction of a sufficiently small number of infected individuals does not lead to disease outbreak in the population. In other words, the system undergoes a bifurcation whereby the stability of the DFE changes as $\mathcal{R}_0$ passes through 1.

In our model, V^{-1} may be computed explicitly (but is too unwieldy to reproduce here), and since F has only one non-zero row, the product FV^{-1} has only one non-zero row. Therefore FV^{-1} has only a single non-zero eigenvalue, whose modulus is the spectral radius. Since the associated eigenvector is $(1, 0, \ldots, 0)$, the non-zero eigenvalue is given simply by the first matrix entry, $(FV^{-1})_{11}$, that is, the product of the first row of F with the first column of V^{-1}, namely

$$\mathcal{R}_0 = (1-u_s)S^*\left(R_A + R_P + R_I\right), \tag{42a}$$

$$R_A = \frac{\beta_A}{\gamma_A}\frac{1}{\Lambda}\Big\{p_a\Big[(\gamma_I + \tau(0))(C_{A,P} + C_{P,P} + \gamma_P) + (C_{I,I} + C_{A,I} + C_{P,I})\gamma_P + C_{I,I}C_{P,P} - C_{I,P}C_{P,I}\Big] + (C_{A,I}C_{I,P} - C_{A,P}C_{I,I})(1-p_a) - C_{A,P}(\gamma_I + \tau(0)) - C_{A,I}\gamma_P\Big\}, \tag{42b}$$

$$R_P = (1-p_a)\beta_P\frac{\gamma_I + \tau(0) + C_{I,I}}{\Lambda}, \tag{42c}$$

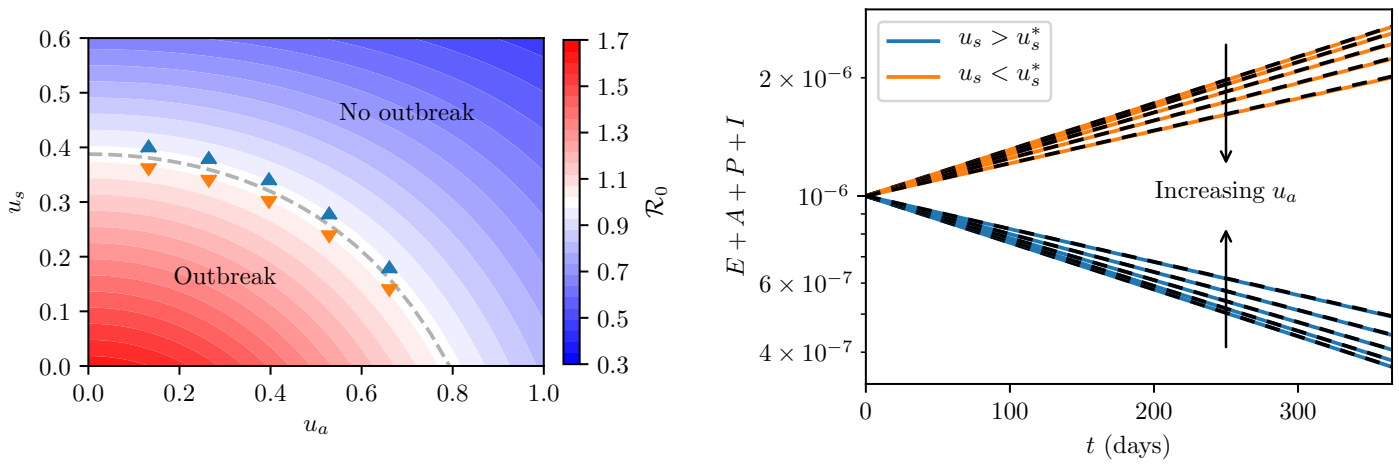
$$R_I = (1-p_a)\beta_I\frac{\gamma_P - C_{I,P}}{\Lambda}, \tag{42d}$$

$$\Lambda = [(1-p_a)C_{E,I} + C_{P,I}](\gamma_P - C_{I,P}) + [(1-p_a)C_{E,P} + C_{P,P} + \gamma_P](\gamma_I + \tau(0) + C_{I,I}). \tag{42e}$$

In the absence of contact tracing, $u_a = 0$, the contact tracing terms $C_{K,J}$ vanish and we find that

$$\mathcal{R}_0|_{u_a=0} = (1-u_s)S^*\left[p_a\frac{\beta_A}{\gamma_A} + (1-p_a)\left(\frac{\beta_P}{\gamma_P} + \frac{\beta_I}{\gamma_I + \tau(0)}\right)\right]. \tag{43}$$

In Figure 4, we demonstrate that the $\mathcal{R}_0$ calculation produces a threshold parameter by comparing numerical solutions to the asymptotic predictions on either side of the $\mathcal{R}_0 = 1$ level set. The excellent agreement suggests that the analysis that provides a way to evaluate at the DFE is representative of the system dynamics. That all the parameters below the $\mathcal{R}_0 = 1$ level set exhibit disease outbreak, while all those above it exhibit no outbreak validates with the theoretical predictions.



(a) (b)

Figure 4: ***(a)*** *Basic reproduction number $\mathcal{R}_0$ (duplicate of Figure 3 left of the main paper). The dashed line indicates the critical level set $\mathcal{R}_0 = 1$. Marks on either side of the critical level set correspond to parameters used in (b) for comparison.* ***(b)*** *Total infections $E + A + P + I$ over time for different u_a and u_s values marked in (a).*

S3 Further numerical simulations

S3 A Outbreak prevention

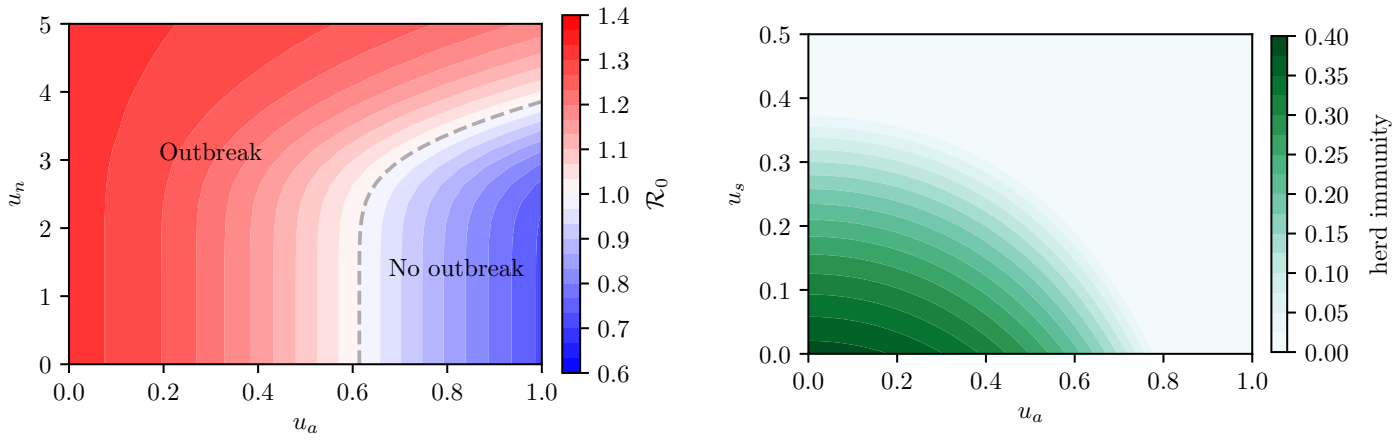
In this section, we study further how the basic reproduction number sheds light on outbreak prevention. In Figure 5a we plot the basic reproduction number $\mathcal{R}_0$ in (u_a, u_n)-space. We see that there is little variation in $\mathcal{R}_0$ for $u_n \lessapprox e_{1/2} - \mathcal{O}(e_0)$. In Figure 5b we illustrate, in (u_a, u_s)-space, the herd immunity required $1 - S^*$ to ensure no outbreak will occur. This quantity is particularly relevant after a wave of the disease has passed through the population. If social intervention measures were put in place, relaxing these might result in a second disease outbreak if there is insufficient herd immunity. From another perspective, Figure 5b demonstrates how relaxing intervention measures may be justified as a function of herd immunity.

In Figure 6 we plot intervention thresholds (level sets of parameter space where $\mathcal{R}_0 = 1$, separating the space into regions where a disease outbreak does or does not occur) when the entire population is initially susceptible $S^* = 1$, to explore how these change with notification thresholds u_n, social intervention levels u_s, and testing rates τ_0. In Figure 6a we see that, as we increase the notification threshold u_n, the effect of contact tracing is diminished as infected contacts are not being informed to isolate. In Figure 6b we observe that, with increased testing rates τ_0, the outbreak region of parameter space shrinks. This is because faster testing results in earlier isolation of infectious individuals, as well as increased contact tracing. At low adoption rates $u_a \approx 0$, the threshold is improved only due to testing, while at larger u_a the gains are amplified. Similarly, in Figures 6c and 6d increasing social intervention u_s and testing rates τ_0 results in a smaller outbreak region of parameter space.

These additional results exhibit the importance of the basic reproduction number in determining the conditions to prevent a primary or secondary outbreak in the high-dimensional parameter space.

S3 B Optimal notification threshold

In this section, we further demonstrate the nature of the trade-off inherent in choosing a notification threshold. In Figure 7b we plot the proportion of infections in the population as a function of time for various notification thresholds u_n. We see that as the notification threshold is lowered, the more aggressive contact tracing is more effective in eradicating the disease. However, In Figure 7a we plot the corresponding quarantine costs. We see that for u_n below the optimal, the quarantine costs are inflated. As u_n is increased beyond the optimal value, the quarantine costs are initially lower, because the tracing is more precise. However, over time the cost accumulates because the disease is not being eradicated effectively.



(a) (b)

Figure 5: ***(a)*** *Basic reproduction number $\mathcal{R}_0$ in (u_a, u_n)-space for $u_s = 0.3$. The dashed line indicates the critical level set $\mathcal{R}_0 = 1$, which is the intervention threshold separating the parameter space where there will or will not be an outbreak.* ***(b)*** *Herd immunity $1 - S^*$ required to prevent an epidemic. Beyond the $\mathcal{R}_0 = 1$ level set no herd immunity is required. However, if social interventions and contact tracing adoption are such that $\mathcal{R}_0 > 1$, then a proportion of the population needs to have obtained immunity to prevent disease outbreak.*

S3 C Infections saved by new adopters

In this section we calculate an expression approximating the *number* of new adopters required to reduce the total number of infections by one. $S + Q$ quantifies the uninfected (susceptible and healthy quarantined) proportion of the population. We emphasise the dependence of the dynamics on the contact tracing adoption fraction u_a by writing $(S+Q)(t; u_a)$, and denote the population size by N. The *number* of infections saved over a time horizon T due to an additional adopter may then be written as

$$N(1 - (S+Q)(T; u_a)) - N(1 - (S+Q)(T; u_a + 1/N)) = \frac{\partial (S+Q)}{\partial u_a}(T; u_a) + \mathcal{O}(1/N) \sim \frac{\partial (S+Q)}{\partial u_a}. \tag{44}$$

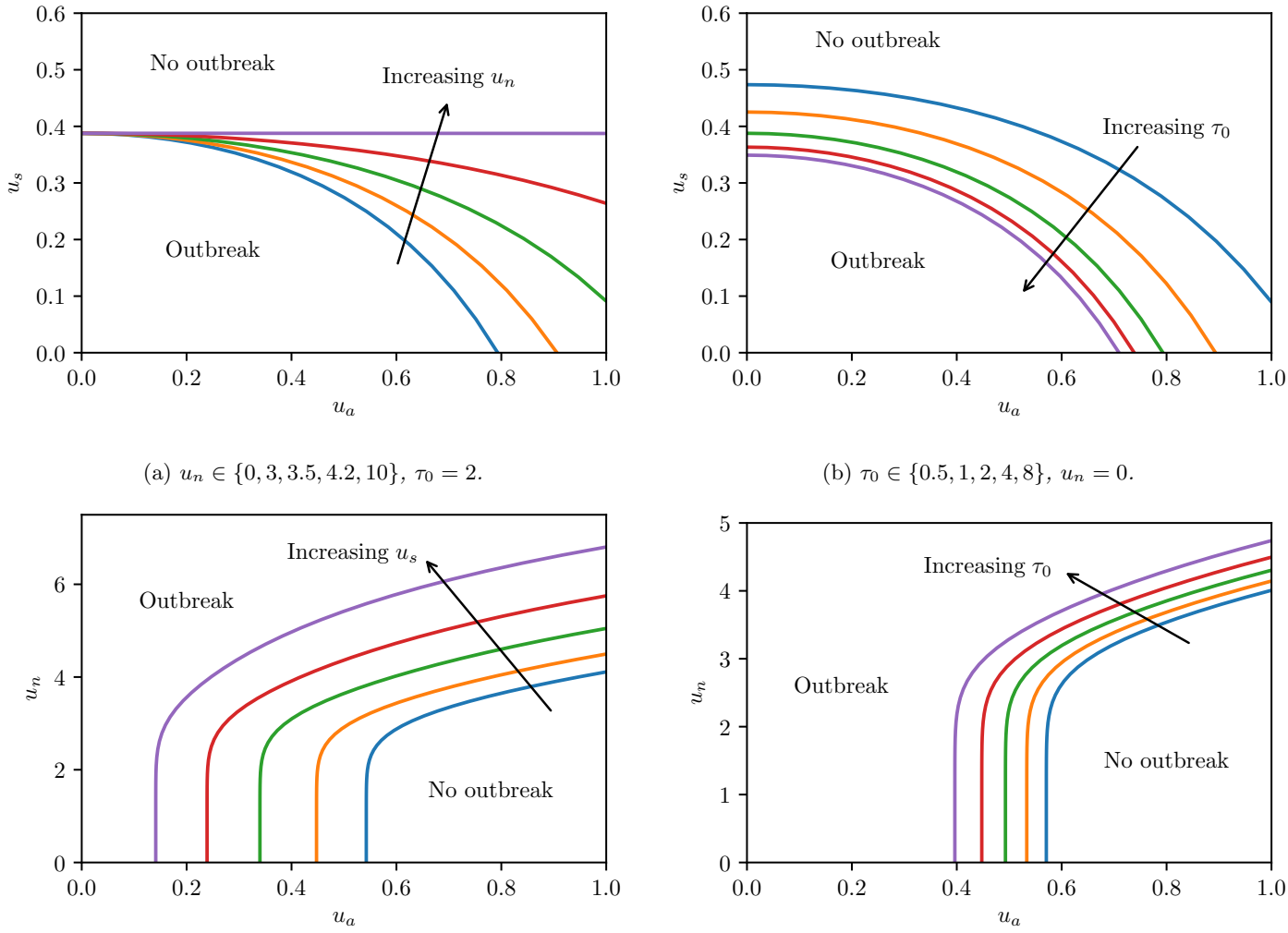
The inverse of (44) is the quantity of interest, representing the number of adopters required to save an infection.

The inverse of (44) is plotted in Figure 8 for $T = 3$ years. For small values of u_a and u_s, additional contact tracing participation helps reduce infections. However, when u_a is too low, the influence of additional adopters is dramatically reduced (reflecting the dependence on u_a^2 versus u_s). In these cases, the epidemic has passed through the population within three years.

Near the $\mathcal{R}_0 = 1$ boundary, additional adopters save many infections by postponing the disease peak beyond the time horizon. Thus, sufficiently close to the boundary where the disease does not significantly invade within three years, new adopters make less impact. Ultimately, in both of these two regions, where $[\partial(S+Q)/\partial u_a]^{-1} \leq 1$, just a single new adopter can save another from infection.

S3 D The impact of tracing delay

In this section we explore how a delay t_0 in contact tracing influences its efficacy. Recall from Section S3 D that a traced contact is calculated to have progressed to compartment $K \in \{E, A, P, I, R\}$ by the time of tracing with probability μ_K. A delay of t_0 days in the contact tracing manifests as a redistribution of the traced contacts among the compartments via μ_K. For small t_0, most of the traced contacts will be in compartment E, but as t_0 increases, traced contacts are more likely to have been removed (R), as shown in Figure 9a. With increasing t_0, as the probability of traced contacts being removed rises, the efficacy of contact tracing diminishes (Figure 9b). We note that Figure 9b is strikingly similar to Figure 6a: in both the limit $t_0 \to \infty$ and the limit $u_n \to \infty$ contact tracing has no effect, therefore, these two parameters must interpolate between the same two limiting cases.



(a) $u_n \in \{0, 3, 3.5, 4.2, 10\}$, $\tau_0 = 2$.

(b) $\tau_0 \in \{0.5, 1, 2, 4, 8\}$, $u_n = 0$.

(c) $u_s \in \{0.25, 0.3, 0.34, 0.365, 0.38\}$, $\tau_0 = 2$.

(d) $\tau_0 \in \{1, 1.2, 1.5, 2, 3\}$, $u_s = 0.3$.

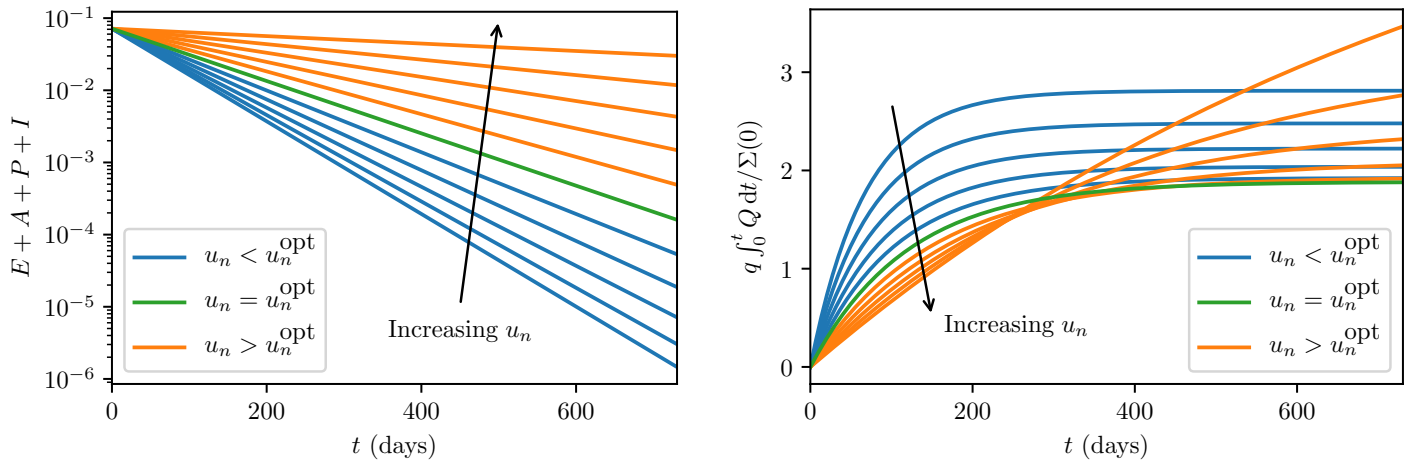
Figure 6: *Intervention threshold determining whether there will be disease outbreak for various social intervention levels u_s notification thresholds u_n and testing rates τ_0 in **(a,b)** (u_a, u_s)-space; **(c,d)** (u_a, u_n)-space.*

S3 E Robustness to deviations in the distributions

In this section, we demonstrate how our results change with changes in the distributions ρ_c and p_i by varying the infection uncertainty e_0, the infectious exposure level $e_{1/2}$, and the contact variability c_0. To do this, we explore the intervention threshold curve (the level set $\mathcal{R}_0 = 1$) in (u_a, u_n)-space (as illustrated in Figures 5a, 6c and 6d).

It is instructive to distinguish between two sections of the intervention threshold curve. For small u_n the curve is nearly vertical: for increasing u_n the increases in u_a are exponentially small. This is a consequence of low-exposure contacts having an exponentially small probability of infection, therefore, increasing u_n (corresponding to not notifying contacts who are unlikely infected) has a negligible effect on the contact tracing precision. When $u_n \approx e_{1/2} - \mathcal{O}(e_0)$ the curve turns, and with increasing u_n, increases in u_a are of the same magnitude. In this section of the curve, contacts are all very likely to have been infected, therefore, to maintain the intervention threshold with increasing u_n requires additional contact tracing adoption u_a. We can derive a simple form that approximates the curvature in this section of the curve by noting that $p_i \approx 1$ (up to an exponentially small correction). Therefore, the contact tracing term $\alpha\Theta$ given in (40) satisfies

$$\begin{aligned} \alpha\Theta &\propto u_a^2 \int_{u_n}^{\infty} \rho_c(e) p_i(e)\, \mathrm{d}e \\ &\approx u_a^2 f_c(u_n), \end{aligned} \tag{45}$$



(a) (b)

Figure 7: ***(a)*** *Infectious proportion;* ***(b)*** *quarantine cost; versus time for $u_a = 0.5$, $u_s = 0.35$, and different notification thresholds u_n. Denoting the optimal notification threshold by u_n^{opt} and the critical notification threshold (for which $\mathcal{R}_0 = 1$) by u_n^*, the simulations use values $u_n = u_n^{\text{opt}} + (u_n^* - u_n^{\text{opt}})i/5$ for $i = -5, -4, \ldots, 5$. The simulations correspond to Figure 4 left of the main paper.*

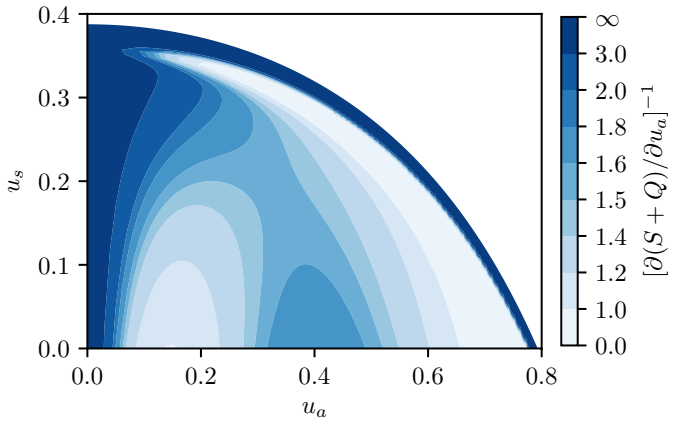


Figure 8: *The number of new contact tracing adopters required to save one infection over a three-year period. There are two primary regions where this quantity is low (i.e. each new contact tracing adopter saves a relatively large number of infections): low social intervention measures u_s with moderately low adoption fraction u_a, and also close to the $\mathcal{R}_0 = 1$ boundary. In these regions, each new adopter saves another from infection.*

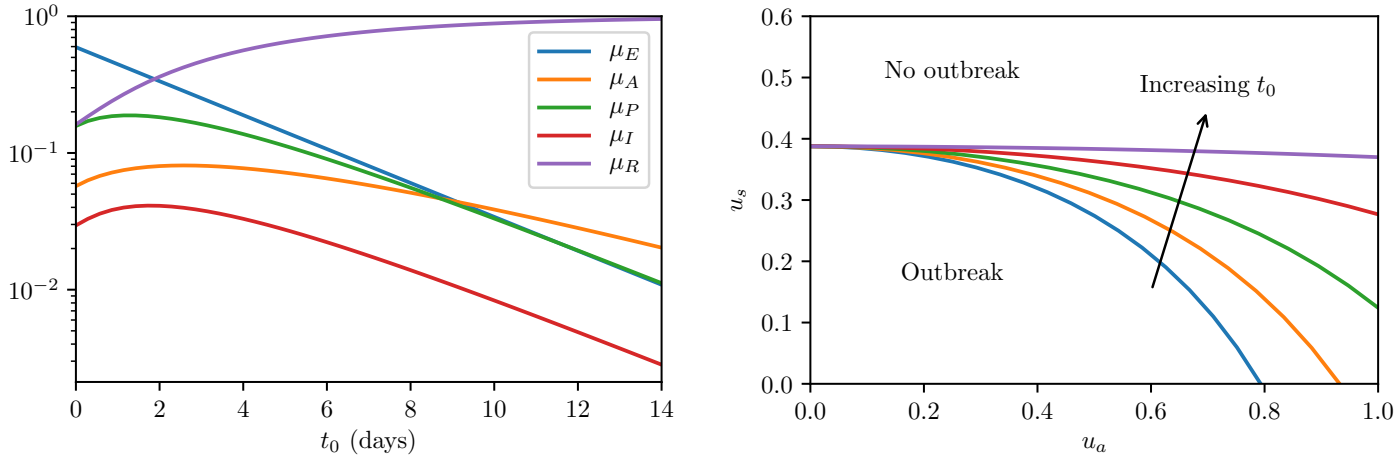
where the constant of proportionality is independent of u_a and u_n. With the parametrisation (12), this takes the form

$$\alpha\Theta \propto u_a^2 \, \mathrm{e}^{-u_n/c_0}. \tag{46}$$

The intervention threshold is given by the level set $\mathcal{R}_0 = 1$, on which $\alpha\Theta$ is constant when all else is fixed. It then follows that $\mathrm{d}u_n/\mathrm{d}u_a \approx 2c_0/u_a$, and thus,

$$u_n \approx 2c_0 \log(u_a) + \text{const.} \tag{47}$$

The approximation (47) is plotted as a black dashed curve in Figure 10a, where we find excellent agreement with the exact solution. We thus call this curve section the logarithmic segment.



(a) (b)

Figure 9: ***(a)*** *Probabilities μ_K for $K \in \{E, A, P, I, R\}$ as a function of the tracing delay t_0.* ***(b)*** *Intervention threshold determining whether there will be disease outbreak for tracing delays $t_0 \in \{0, 2, 4, 7, 14\}$.*

Increasing the infection uncertainty e_0 leads to requiring more intervention (either higher u_a for fixed u_n or lower u_n for fixed u_a) to prevent an outbreak. While the general structure of the intervention threshold curve is preserved, the transition between the vertical and the logarithmic segments becomes less sharp with increasing infection uncertainty.

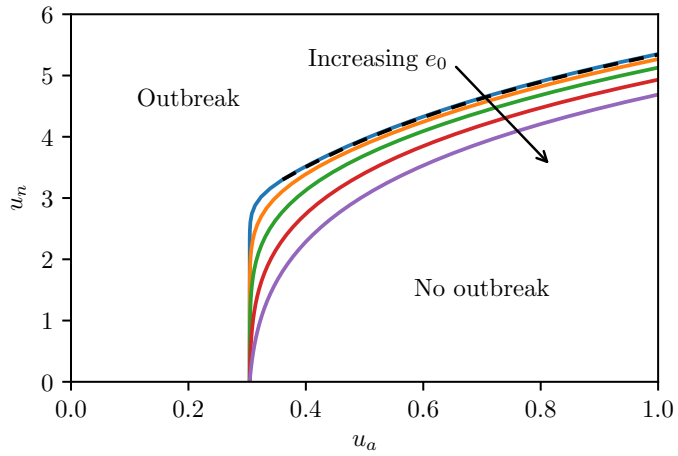
Increasing the infectious exposure level $e_{1/2}$ serves to shift the intervention threshold curve (Figure 10b). This is the result of a change in $e_{1/2}$ corresponding to a translation in p_i (see (12)). It then follows from (40) that, under the transformation $e_{1/2} \mapsto e_{1/2} + x$, the contact tracing term $\alpha\Theta$ is transformed via

$$\begin{aligned} \alpha\Theta(e_{1/2}) \mapsto \alpha\Theta(e_{1/2} + x) &\propto u_a^2 \frac{\int_{u_n}^{\infty} \rho_c(e) p_i(e + x)\,\mathrm{d}e}{\int_0^{\infty} \rho_c(e) p_i(e + x)\,\mathrm{d}e} \\ &= u_a^2 \frac{\int_{u_n + x}^{\infty} \rho_c(e - x) p_i(e)\,\mathrm{d}e}{\int_x^{\infty} \rho_c(e - x) p_i(e)\,\mathrm{d}e} \\ &= u_a^2 \frac{\int_{u_n + x}^{\infty} \rho_c(e) p_i(e)\,\mathrm{d}e}{\int_x^{\infty} \rho_c(e) p_i(e)\,\mathrm{d}e} \\ &= u_a^2 \frac{f_i(u_n + x)}{f_i(x)}, \end{aligned} \tag{48}$$

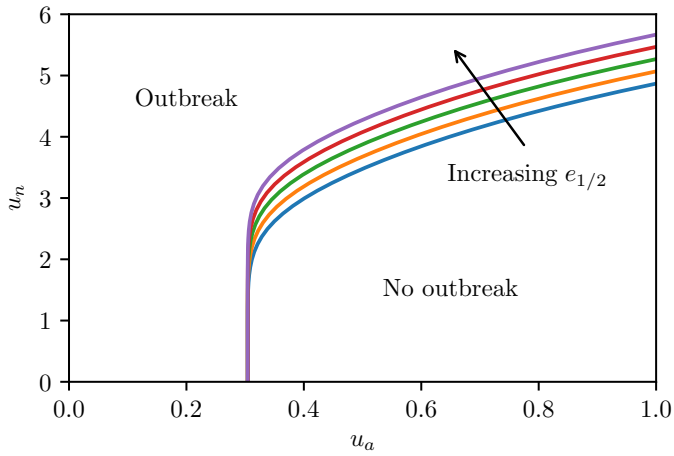
where the constant of proportionality is independent of u_a, u_n, and $e_{1/2}$. The quantity $f_i(x)$ is approximately constant for small x, since changing x corresponds to including/excluding the contribution of low-exposure, hence low-risk, contacts. Thus, the transformation in $\alpha\Theta$ is well approximated by a translation in u_n.

Changing the contact variability c_0 has negligible effect on the vertical segment of the intervention threshold curve, while the slope of the logarithmic segment steepens with increasing c_0 (Figure 10c) as predicted by the approximation (47), which shows excellent agreement with the exact solution. A steeper slope corresponds to more relaxed intervention measures (larger u_n for fixed u_a, or larger u_a for fixed u_e). Increased contact variability results in more high-risk contacts, and hence the contact tracing precision improves, allowing for more relaxed intervention to prevent disease outbreak.

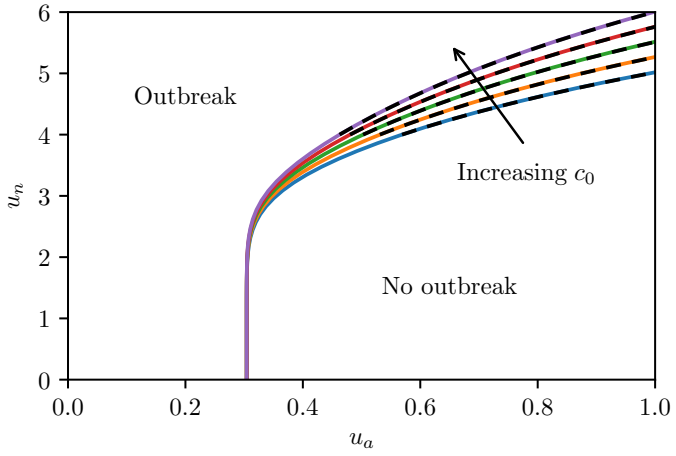
We conclude that, despite quantitative changes, the results remain qualitatively similar when the contact and infection distributions are altered. It follows that the model is robust to uncertainties and small errors in the parametrisation of these distributions.



(a) $e_0 \in \{1/3, 2/3, 1, 4/3, 5/3\}$.



(b) $e_{1/2} \in \{2.6, 2.8, 3, 3.2, 3.4\}$.



(c) $c_0 \in \{0.9, 1, 1.1, 1.2, 1.3\}$.

Figure 10: *Intervention threshold determining whether there will be disease outbreak for social intervention measures* $u_s = 0.35$ *and various distribution parameters* e_0, $e_{1/2}$, *and* c_0. *The black dashed curve illustrates the approximation given in* (47) *where the constant is fixed by choosing a single point on the corresponding solid curve.*

Appendices

A Transition probabilities

In this section, we outline the calculation of the transition probabilities $P_{B|A}$ defined in Section S1 B as

$$P_{B|A}(s) = \mathbb{P}(\text{individual in } B \text{ at } t = s \mid \text{individual in } A \text{ at } t = 0). \tag{49}$$

For the sake of simplifying the notation, we denote the transition rate from compartment K by γ_K. However, this ought to be understood as all contributions to the transition, and is therefore replaced by $\gamma_K + \alpha(1-\Theta)I$ for all $K \neq I$ and $\gamma_I + \tau(I) + \alpha(1-\Theta)I$ for $K = I$.

Starting and finishing in the same compartment is calculated via

$$P_{E|E}(s) = 1 - \mathbb{P}(\text{leave } E \text{ by } t = s) = 1 - \gamma_E \int_0^s \mathrm{e}^{-\gamma_E r}\,\mathrm{d}r = \mathrm{e}^{-\gamma_E s}. \tag{50}$$

Each transition requires a new integration, for example,

$$\begin{aligned} P_{A|E}(s) &= p_a \int_0^s \mathbb{Q}(\text{leave } E \text{ at } t=r)(1-\mathbb{P}(\text{leave } A \text{ by } t=s \mid \text{enter } A \text{ at } t=r))\,\mathrm{d}r \\ &= p_a \int_0^s \gamma_E \mathrm{e}^{-\gamma_E r}\mathrm{e}^{-\gamma_A(s-r)}\,\mathrm{d}r = p_a\gamma_E \frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_A s}}{\gamma_A-\gamma_E}, \end{aligned} \tag{51}$$

One may compute all the transitions in this way, remembering that from compartment E to compartment R one may arrive via two different routes. We find that

$$\begin{aligned} P_{P|E}(s) &= (1-p_a)\gamma_E \frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_P s}}{\gamma_P-\gamma_E}, \\ P_{I|E}(s) &= (1-p_a)\frac{\gamma_E\gamma_P}{\gamma_P-\gamma_I}\left(\frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_P s}}{\gamma_E-\gamma_P} - \frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_I s}}{\gamma_E-\gamma_I}\right), \\ P_{R|E}(s) &= (1-p_a)\left[1-\mathrm{e}^{-\gamma_E s}+\frac{\gamma_E}{\gamma_I-\gamma_P}\left(\gamma_I\frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_P s}}{\gamma_E-\gamma_P} - \gamma_P\frac{\mathrm{e}^{-\gamma_E s}-\mathrm{e}^{-\gamma_I s}}{\gamma_E-\gamma_I}\right)\right] \\ &\quad + p_a\left(1-\mathrm{e}^{-\gamma_E s}+\gamma_E\frac{\mathrm{e}^{-\gamma_A s}-\mathrm{e}^{-\gamma_E s}}{\gamma_A-\gamma_E}\right), \\ P_{I|P}(s) &= \gamma_P\frac{\mathrm{e}^{-\gamma_P s}-\mathrm{e}^{-\gamma_I s}}{\gamma_I-\gamma_P}, \\ P_{I|I}(s) &= \mathrm{e}^{-\gamma_I s}. \end{aligned} \tag{52}$$

When $t_0 = 0$ these take the forms

$$\begin{aligned} p_{E,P} &= \frac{\gamma_P}{(\gamma_E+\gamma_I)(\gamma_E+\gamma_P)}, & p_{E,I} &= \frac{1}{\gamma_E+\gamma_I}, \\ p_{A,P} &= \frac{\gamma_E\gamma_P(\gamma_A+\gamma_E+\gamma_I+\gamma_P)p_a}{(\gamma_A+\gamma_I)(\gamma_E+\gamma_I)(\gamma_A+\gamma_P)(\gamma_E+\gamma_P)}, & p_{A,I} &= \frac{\gamma_E p_a}{(\gamma_A+\gamma_I)(\gamma_E+\gamma_I)}, \\ p_{P,P} &= \frac{\gamma_E(\gamma_E+\gamma_I+2\gamma_P)(1-p_a)}{2(\gamma_E+\gamma_I)(\gamma_E+\gamma_P)(\gamma_I+\gamma_P)}, & p_{P,I} &= \frac{\gamma_E(1-p_a)}{(\gamma_E+\gamma_I)(\gamma_I+\gamma_P)}, \\ p_{I,P} &= \frac{\gamma_E\gamma_P(\gamma_E+\gamma_I+\gamma_P)(1-p_a)}{2\gamma_I(\gamma_E+\gamma_I)(\gamma_E+\gamma_P)(\gamma_I+\gamma_P)}, & p_{I,I} &= \frac{\gamma_E\gamma_P(1-p_a)}{2\gamma_I(\gamma_E+\gamma_I)(\gamma_I+\gamma_P))}, \\ p_{R,P} &= (1-p_a)p_{R,P,0} + p_a\, p_{R,P,1}, & p_{R,I} &= \frac{\gamma_E}{2\gamma_I(\gamma_E+\gamma_I)}\left(\frac{\gamma_P(1-p_a)}{\gamma_I+\gamma_P}+\frac{2\gamma_A p_a}{\gamma_A+\gamma_I}\right), \\ p_{R,P,0} &= \frac{\gamma_E(\gamma_I^2+\gamma_I\gamma_P+\gamma_E\gamma_I+\gamma_E\gamma_P+\gamma_P^2)}{2\gamma_I(\gamma_E+\gamma_I)(\gamma_E+\gamma_P)(\gamma_I+\gamma_P)}, & p_{R,P,1} &= \frac{\gamma_A\gamma_E(\gamma_I^2+\gamma_A\gamma_E+\gamma_I\gamma_P+\gamma_P^2+(\gamma_A+\gamma_E)(\gamma_I+\gamma_P))}{\gamma_I(\gamma_A+\gamma_I)(\gamma_E+\gamma_I)(\gamma_A+\gamma_P)(\gamma_E+\gamma_P)}. \end{aligned} \tag{53}$$

References

[1] Y. M. Bar-On, A. Flamholz, R. Phillips, and R. Milo. Science forum: Sars-cov-2 (covid-19) by the numbers. *eLife*, 9:e57309, mar 2020.

[2] M. Begon, M. Bennett, R. G. Bowers, N. P. French, S. M. Hazel, and J. Turner. A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiology and Infection*, 129(1):147–153, 2002.

[3] C. Browne, H. Gulbudak, and G. Webb. Modeling contact tracing in outbreaks with application to ebola. *Journal of Theoretical Biology*, 384:33–49, 2015.

[4] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio r0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[5] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser. Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science*, 2020.

[6] J. K. Hale. *Ordinary Differential Equations*. Dover Books on Mathematics Series. Kreiger Publishing Company, 2 edition, 1980.

[7] E. J. Hinch. *Perturbation Methods*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 1991.

[8] J. M. Hyman, J. Li, and E. A. Stanley. Modeling the impact of random screening and contact tracing in reducing the spread of HIV. *Mathematical Biosciences*, 181(1):17–54, 2003.

[9] S. Lee, R. Morales, and C. Castillo-Chavez. A note on the use of influenza vaccination strategies when supply is limited. *Mathematical Biosciences and Engineering*, 8(1):171–182, 2011.

[10] K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*, 25(10), 2020.

[11] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1):29–48, 2002.

[12] W. Wang. Backward bifurcation of an epidemic model with treatment. *Mathematical biosciences*, 201(1–2):58–71, 2006.