

Counter-intuitive COVID-19 Trajectories – Explanations, Early Warning Indicator and Mitigation Strategies

B SHAYAK (*)

Theoretical and Applied Mechanics,
Mechanical and Aerospace Engg,
Cornell University,
Ithaca – 14853,
New York State, USA

MOHIT M SHARMA

Population Health Sciences,
Weill Cornell Medicine,
1300 York Avenue,
NYC – 10065,
New York State, USA

(*) Corresponding author. Email : sb2344@cornell.edu , shayak.2015@iitkalumni.org
ORCID : 0000-0003-2502-2268

Thursday 16 September 2021

---- O ----

Keywords. Cluster seeding and transmission model Contact tracing Sensitivity to parameters and initial conditions Tracing capacity ratio

Abstract. The COVID-19 trajectories worldwide have shown several surprising features which are outside the purview of classical epidemiological models. These include (a) almost constant and low daily case rates over extended periods of time, (b) sudden waves emerging from the above solution despite no or minimal change in the level of non-pharmaceutical interventions (NPI), and (c) reduction or flattening of case counts following relaxation of NPI. To explain these phenomena, we add contact tracing to our recently developed cluster seeding and transmission (CST) model. We find no fewer than four effects which make prediction of epidemic trajectories uncertain. These are (a) cryptogenic instability, where a small increase in population-averaged contact rate causes a large increase in cases, (b) critical mass effect, where a wave manifests after weeks of quiescence with no change in parameter values, (c) knife-edge effect, where a small change in parameter across a critical value causes a huge change in the response of the system, and (d) hysteresis effect, where the timing and not just the strength of a particular NPI determines the subsequent behaviour. Despite these effects however, some non-obvious conclusions regarding NPI appear to be robust. In particular, (a) narrowing the circle of one's social interactions can be as effective a measure as reducing interactions altogether, and (b) a good contact tracing program can effectively substitute for much more invasive measures. Finally, we propose the contact tracing capacity ratio – a metric of the load to which the tracers are subject – as a reliable early warning indicator of an imminent epidemic wave.

Word counts. Abstract – 256, Main Article text – 9477, Grand total – 11502

---- O ---- O ---- O ---- ---- O ---- O ---- O ----

INTRODUCTION

§1. COVID-19 curves and existing model predictions. The dynamics of COVID-19 trajectories in different parts of the world have shown many features which are unexpected from a traditional epidemiological viewpoint. “Plateaus” or extended periods of low and nearly constant daily case rate have been seen for example in USA (June 2021), UK (April to May 2021), India (January to March 2021), Canada (June to August 2020), Germany (May to August 2020), Uruguay (March to November 2020) and Taiwan (March 2020 to April 2021). In each case, the quiescent period was shattered by a sudden wave; in at least some instances, for example India and Taiwan, non-pharmaceutical interventions (NPI) had already been significantly relaxed long before the wave arose and there was no change in public policy immediately preceding the wave. In UK, a wave ostensibly driven by the extreme transmissibility of the B1.617.2 double mutant (“delta variant”) began in late May 2021; despite this, the government went ahead with “Freedom Day” (relaxation of all restrictions) on 19 July 2021, after which the case counts went down. This phenomenon was described as “Nobody really knows what’s going on” [1] by a senior scientist in a research group which has published dozens of scientific papers on COVID-19 mathematical modeling. In India also, following the horrific second wave, B1.617.2 (the dominant strain) remains under control in almost all states. However, in parts of USA and Israel, despite high vaccination coverage, the case counts are going from bad to worse with the double mutant being the dominant strain and shouldering almost all of the blame.

The prevalence of different mutants with time in India [2] and USA [3] is also interesting. For example, in India, the prevalence of B1.617.2 increased rapidly between March and May (the time-frame of the wave), as is expected for a highly transmissible strain. The surprise however is that the wildtype virus B1 held on to a 10-15 percent market share while the allegedly more transmissible [4] B1.1.7 (“alpha”) mutant became reduced to nearly zero prevalence. In USA too, B1.617.2 has proliferated rapidly between May and today; once again wildtype has clung to a steady non-zero prevalence while B1.1.7 has been all but knocked out of the competition.

In Ref. [5] we have proposed a lumped parameter or compartmental model based on delay differential equation (DDE) which we believe is superior to other existing compartmental models, for reasons elaborated therein. Even this model however cannot explain any of the above ‘anomalies’. A plateau is a non-generic solution and, for a given infection level, occurs only in a set of measure zero in the parameter space. Sudden waves can only be generated by sudden changes in parameters, which correspond to sudden lifting of NPI. While some of the waves were directly linked to relaxations or COVID-inappropriate behaviour (for example, the massive third wave in USA was linked to travel and festivities during the Thanksgiving-Christmas-New Year season), others, as we have already mentioned, occurred without any such obvious trigger. This latter is completely outside the purview of lumped parameter models. Behaviour of mutants in compartmental models is very simple – a more transmissible mutant always dominates a less transmissible one, irrespective of whether cases are increasing or decreasing.

§2. Cluster seeding and transmission (CST) model – prior work. To address the issue of the plateau solution and the sudden waves, we have proposed a novel cluster-based mathematical model of disease transmission in Ref. [6] and used it in Ref. [7] to perform a detailed analysis of the origins of the devastating second wave in India. These two studies together with Ref. [5] also summarize the state of the art in mathematical modeling of COVID-19 and we refer to these for a detailed review of literature which can set the present Article in a more extensive context. An unstated or at best understated assumption at the heart of all lumped parameter models is that of homogeneous mixing – the assumption that any random person has equal probability of interacting with, and hence transmitting the virus to, any other random person taken from the entire population. In our cluster-based model we have gone with the more realistic assumption that a person is more likely (perhaps by order of magnitude) to interact with (and consequently infect) family members and friends than strangers. We have expressed this by dividing the population into clusters or small groups of people with dense links among each other and few or no links outside. Intra-cluster transmission of virus is certain and rapid. Inter-cluster transmission is rarer and (as per the model assumptions) can occur in two ways. These are unintentional cluster transition or UCT events which

involve accidental transmission in public places, and socializing external to cluster or SEC events where people from different clusters intentionally get together. By calculating the various transmission probabilities, we have arrived at a computational model for the spread of the disease.

We now give a brief description of CST model. Time is discrete and is measured in days. The variables and parameters are as follows (Table adapted from Ref. [6]). The default values are those used in Ref. [6] to generate a plateau solution, and carry over to this Article with a few exceptions to be mentioned as appropriate.

Variable	Significance
y_i	Cumulative number of socially active cases upto and excluding day # i
Δy_i	Number of new socially active cases cropping up on day # i
z_i	Cumulative number of insusceptible clusters upto and excluding day # i
Δz_i	Number of new clusters turning insusceptible on day # i
f_i	Cumulative number of cautious household cases upto and excluding day # i
Δf_i	Number of new cautious household cases cropping up on day # i

Parameter	Significance	Default value
N	Total population	3,02,400
h	Household size	3
N_1	Socially active population	1,00,800
N_C	Number of clusters	4200
s	Size of the cluster	24
\mathbf{v}	Cluster sequence	[1; 3; 6; 7; 5; 1]
T_s	Serial interval	5 days
n_U	Number of people participating daily in UCT events	10,000
P_U	Probability of transmission in a UCT event	0.15
n_S	Number of people participating daily in SEC events	—
m_S	Number of people infected by one case at a SEC event	2
—	Time between exposure and start of transmissibility for all cases	5 days
—	Duration for which all cases transmit to others	3 days

Table 1 : Variables and parameters in the basic CST model. Large parts of this Table are reproduced verbatim from Ref. [6]. We do not assign a default value to n_S (it was zero in Ref. [6]) since it will be varied incessantly in the bulk of this Article.

We count a person as a case on the day s/he first turns transmissible. Since we assume homogeneous mixing among clusters (rather than individuals), the domain of validity of the model is a city rather than a state or a country. Considering a city of total N people, we have partitioned them into N_1 ‘socially active’ people who can contract the disease on their own and $N - N_1$ ‘householders’ who can contract the disease only from a socially active individual. We have then divided the N_1 active people into N_C clusters of equal size s . When a cluster is seeded, i.e. when the first member of a cluster with all susceptible persons is exposed to an infective dose of the pathogen, we assume that cases crop up in the cluster over the next few days following a fixed, definite sequence. This sequence approximately captures the reproduction number R_0 of the disease (R is the number of secondary transmissions from one infected individual, and R_0 is its initial value when everyone is susceptible) and the serial interval (the time between primary and secondary case). The last two rows of Table 1 indicate our assumption that all cases turn transmissible 5 days following

exposure and remain transmissible and at large for 3 days (after this duration, asymptomatic cases recover while symptomatic ones show their symptoms and go into quarantine). We shall in general continue using the same values here.

In the basic CST model, we have found two non-classical effects which are of considerable interest – the cryptogenic instability where an increase in n_s causes a disproportionate rise in case count relative to the increase in population-averaged interaction rate, and the critical mass effect where, for the same parameter values, a very small initial condition results in a containment solution whereas a slightly larger initial condition causes a huge wave.

§3. Objective of this study. While Ref. [6] went a long way towards our understanding of COVID-19 trajectories, it suffered from an important limitation. This was that the model did not incorporate testing, tracing and quarantining of cases. This is a very important public health measure which can significantly influence the case trajectories. For this reason, Ref. [6] also did not go too much into the details of applying NPI to mitigate the epidemic.

In this work, we address these drawbacks. We formulate the CST model with contact tracing in §4-5, present the novel effects in §6, and discuss them in detail in §7. In §8 we turn to the limitations of the model, highlighting features which are expected to remain robust when the assumptions and approximations are relaxed. In §9 we propose an early warning indicator of an imminent wave, and in §10 we briefly mention the public health implications of our findings. We then wrap things up with a conclusion.

---- 0 ----

MATHEMATICAL MODEL

§4. Implementation of contact tracing. Before starting the discussion, we get one assumption out of the way. This is that our model does not incorporate vaccination. A detailed analysis of this intervention is appropriate for a future, separate study while an approximate estimate of its effect can be obtained simply by considering a fraction (vaccinated percentage times vaccine efficacy) of the population to be pre-immune.

Contact tracing refers to the process in which, when a case is detected, the people with whom s/he has interacted recently are identified and recommended to isolate for a few days so as to prevent further spread of infection. Here we make two fundamental assumptions regarding the tracing process :

- There is forward contact tracing only, starting from symptomatic cases. This means that when a symptomatic case with unknown source of exposure tests positive, the authorities attempt to identify all of his/her potential secondary cases, but do not attempt to identify the source of exposure and other secondary contacts thereof. We also assume that there is no random testing of asymptomatic persons.
- The process timings involved are such as to ensure that cases who are successfully caught by the contact tracers are sent into quarantine during their non-transmissible incubation period, i.e. traced cases transmit the disease to no one else.

A toy example will help to clarify the implications of the above assumptions. For this example we ignore the cluster structure of the population. Let us say Alfa is an asymptomatic case who transmits the virus to Bravo and Charlie on day #0. On day #5, both of them turn transmissible – Bravo transmits to Delta while Charlie to Echo and Foxtrot, all on the same day. Bravo remains asymptomatic while on day #8, Charlie turns symptomatic, reports to the authorities and tests positive. By the first assumption, Echo and Foxtrot are rounded up by the contract tracers but there is no attempt to track down Alfa. As a result of this lapse Bravo remains undetected as well, and so does Delta. Charlie's secondary cases Echo and Foxtrot turn transmissible on day #10; the second assumption implies that they are successfully captured within day

#9. This discussion might appear a little unwieldy so we show the transmission sequence in a schematic form below.

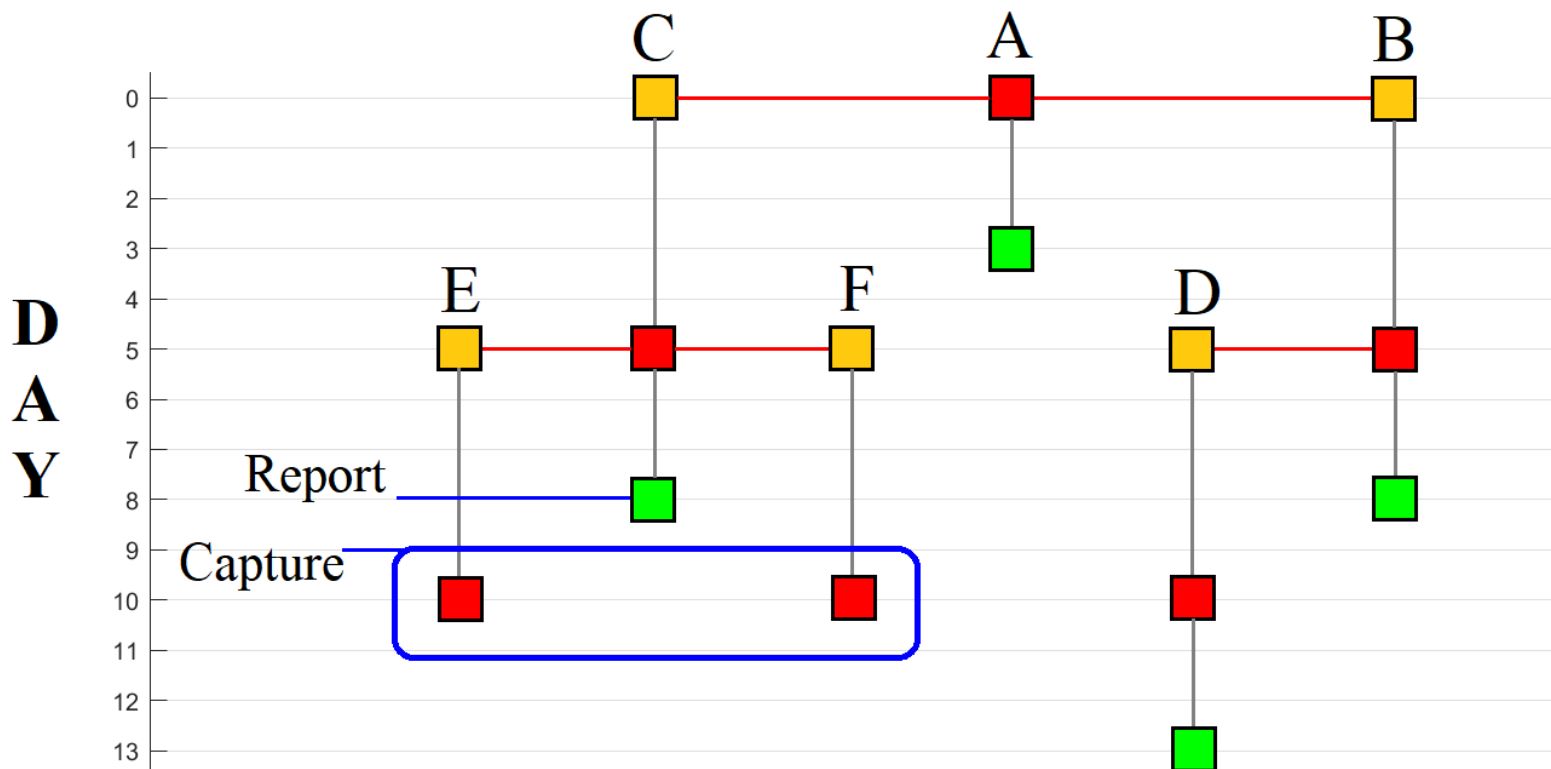


Figure 1 : Timeline of different cases in the hypothetical example to demonstrate the contact tracing assumptions. Squares denote individual persons. Yellow means exposed but not transmissible, red means transmissible while green means recovered. The grey vertical lines indicate transition of one person from one state to another while red horizontal lines denote transmission of the disease from person to person. The 5- and 3-day intervals between exposure and infectiousness, and infectiousness and recovery are consistent with the basic CST model assumptions. Note how the two assumptions regarding contact tracing result in capture of Echo and Foxtrot before they turn transmissible while failing to trace the Bravo-Delta branch of the transmission.

Thus we can see that the two assumptions introduce errors in opposite senses. A more sophisticated contact tracing system can achieve backward tracing (catching Alfa from Charlie and hence Bravo and Delta from Alfa) while a slowly functioning system with delays in testing etc can result in secondary cases not being caught before starting spread.

The second assumption allows us to define exactly two classes of cases – at large cases, who spend three full days spreading virus just as in Ref. [6] and quarantined cases who spend zero days infecting others. We now account for the presence of social clusters in a realistic manner. Let P_0 be the probability that a random case is symptomatic, and $Q_0 = 1 - P_0$ the probability that s/he is asymptomatic. We assume that when a symptomatic case reports to the authorities, the entire cluster to which s/he belongs is successfully identified with probability P_1 and missed with probability $1 - P_1$. Further, UCT and SEC transmissions which this case has caused are identified with probabilities P_2 and P_3 respectively.

We use an example rather than a general theoretical discussion to understand how contact tracing will work on a cluster, noting that the example is very easy to generalize. For this example, we use the cluster sequence [1; 3; 6; 8; 5; 1] which shall also feature in very many of the simulation runs. This sequence is adapted from Table 1, with the reason for the change coming below. All clusters are first seeded via UCT or SEC events; again for definiteness, let us focus on the former. Let the example cluster be named Team, let it consist of persons Alfa through Xray (total 24) and let the at large case Yankee be the one who exposes Alfa to the pathogen in a UCT event. In the absence of contact tracing, the cluster sequence implies that

- Alfa (1 case) is exposed by Yankee on day #0 and turns into a case on day #5
- Bravo, Charlie and Delta (3 more cases) are exposed by Alfa and turn into cases on day #10
- Echo through Juliett (6 more cases) are exposed by one or more of Bravo through Delta and turn into cases on day #15

- Kilo through Romeo (8 more cases) are exposed by one or more of Echo through Juliett and turn into cases on day #20
- Sierra through Whiskey (5 more cases) are exposed by one or more of Kilo through Romeo and turn into cases on day #25
- Xray (1 more case) is exposed by one or more of Sierra through Whiskey and turns into a case on day #30

Now incorporate contract tracing on Team. There is a probability P_0 that Yankee is symptomatic and Q_0 that he is asymptomatic. In the latter case, there is no hope of catching Alfa; in the former case, there is a probability P_2 that Alfa is identified by the contact tracers and sent into quarantine. If this happens, then the only case in Team is Alfa, who is quarantined at the get-go. In such a situation, we classify Team as a cluster of Type 1. The probability that Team, and by extension any other cluster, is of Type 1 is thus P_0P_2 .

If Yankee is asymptomatic (probability $1 - P_0$) or if he is symptomatic (P_0) but the tracers fail to identify Alfa as his secondary UCT case ($1 - P_2$), then Alfa becomes an at large case. The probability of this happening is $1 - P_0 + P_0(1 - P_2)$ which is $1 - P_0P_2$. Then, Alfa transmits the virus to Bravo, Charlie and Delta and also participates in UCT and SEC events. Now, consider the case that Alfa is symptomatic (P_0). If yes, the contact tracers get to work and, with probability P_1 , capture the entire cluster Team including Bravo, Charlie and Delta. Assume the complement $1 - P_1$ to denote a tracing error or roadblock where the opportunity to catch Team is irreversibly lost. The tracers also capture Alfa's UCT and SEC transmissions with probabilities P_2 and P_3 . Captures of UCT and SEC transmissions by members of Team are however accounted for while implementing contact tracing on the secondary clusters, just as we have factored in Yankee while doing the calculation for Team. Hence, for analysing Team, they do not require our further consideration. Bravo, Charlie and Delta contract the infection in quarantine and further spread of the disease within Team is halted. In this situation, we call Team a cluster of Type 2. The probability of this occurring is $(1 - P_0P_2)P_0P_1$.

If Alfa is asymptomatic however (Q_0), then Bravo through Delta perforce become at large cases, transmitting the virus to the next level in the cluster i.e. to Echo through Juliett (and also participating in UCT and SEC). Any and all of Bravo through Delta might be symptomatic or asymptomatic – if at least one is symptomatic ($1 - Q_0^3$, the complement of the probability that all three are asymptomatic) then Team is grounded at this stage with probability P_1 . Echo through Juliett become quarantined cases and we classify Team as Type 3. The probability of this occurring is $(1 - P_0P_2)Q_0(1 - Q_0^3)P_1$.

Similarly, if everyone upto Delta is asymptomatic (Q_0^4) but at least one among Echo through Juliett is symptomatic ($1 - Q_0^6$), then all 10 of these become at large cases, while Kilo through Romeo are exposed but quarantined before they turn infectious. We call Team a cluster of Type 4, and the probability of its occurrence is $(1 - P_0P_2)Q_0^4(1 - Q_0^6)P_1$. The probability that everyone from Alfa to Juliett is asymptomatic is minuscule and we take it as zero. Thus, we define the Type 5 cluster to be the one where contact tracing has no effect at all i.e. all cases remain at large. The probability of its occurrence is 1 minus all the above probabilities put together. A picture might well be worth the last 500 or so words, so we present the probability tree below as Fig. 2 and also the case burdens associated with the various cluster types as Table 2. In a similar manner we can account for the probabilities of occurrence of the five types of clusters via SEC events. For this, we replace P_2 by P_3 in the expressions obtained above.

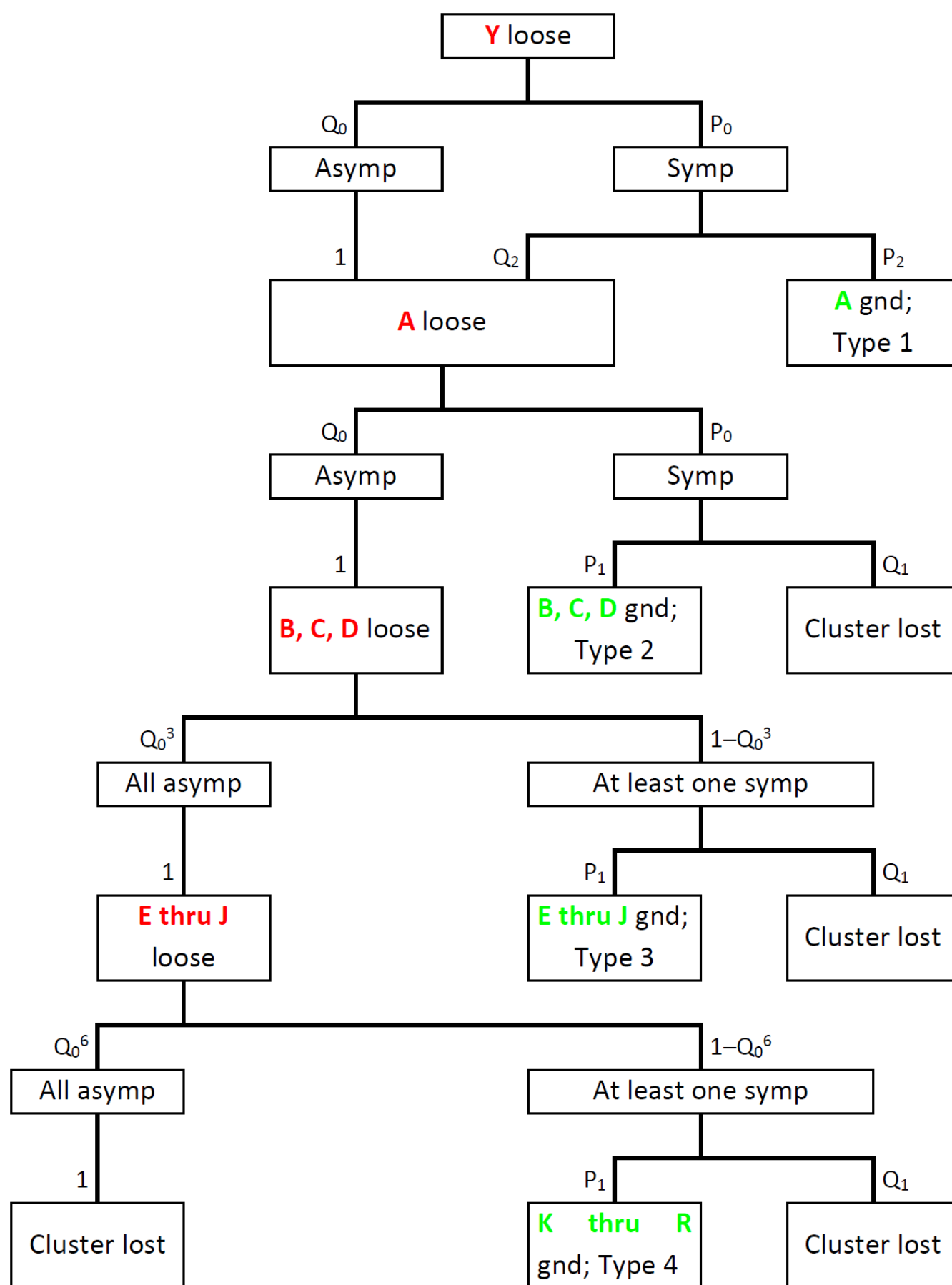


Figure 2 : Probability tree showing how Team can end up being a cluster of Types 1 through 5. “Asymp” and “symp” stand for asymptomatic and symptomatic, “loose” means at large and “gnd” or grounded denotes quarantined. “Cluster lost” means that the tracers fail to track down Team i.e. it becomes a cluster of Type 5. The probabilities associated with each event are shown alongside the vertical arrows leading to the event; Q_n is shorthand for $1 - P_n$ and certainties are marked as 1 for consistency. If deriving (1) from this diagram, note that $Q_0 + P_0 Q_2 = 1 - P_0 P_2$.

Type	1	2	3	4	5
At large	0	1	4	10	24
Quarantined	1	3	6	8	0

Table 2 : Total numbers of at large and quarantined cases arising from clusters of the five fundamental types.

Given these probabilities, we now describe the modifications which must be made to the computational procedure Algorithm 1 of Ref. [6]. Firstly, we define extra variables x_i and Δx_i , the cumulative respectively daily counts of at large cases on day $\#i$ and q_i and Δq_i , the same for quarantined cases. The total number of socially active cases become $y_i = x_i + q_i$ and $\Delta y_i = \Delta x_i + \Delta q_i$. The Subroutine roundoff in Algorithm 1 of Ref. [6], which replaces a fraction by its nearest integer and carries over the error, remains as it is. In the main routine, we redefine the total number of at large cases α present on day $\#i$ to be $\Delta x_{i-1} + \Delta x_{i-2} + \Delta x_{i-3}$. The calculation of expectation number E_U of susceptible clusters seeded via UCT on day $\#i$ remains unchanged from Ref. [6]. This is because contact tracing does not affect the parameters other than α which go into determining E_U . Then, we introduce the probabilities obtained above for the seeded clusters belonging to Types 1 through 5. These probabilities, in summary form, are

$$P_U(1) = P_0 P_2 \quad , \quad (1a)$$

$$P_U(2) = (1 - P_0 P_2) P_0 P_1 \quad , \quad (1b)$$

$$P_U(3) = (1 - P_0 P_2) Q_0 (1 - Q_0^3) P_1 \quad , \quad (1c)$$

$$P_U(4) = (1 - P_0 P_2) Q_0^4 (1 - Q_0^6) P_1 \quad , \quad (1d)$$

$$P_U(5) = 1 - P_U(1) - P_U(2) - P_U(3) - P_U(4) \quad . \quad (1e)$$

Then, the expectation number of susceptible clusters of type j seeded during UCT events becomes $E_U P_U(j)$.

Analogously we calculate the expectation number $E_S(j)$ of susceptible clusters of various types seeded on day $\#i$ during SEC events, by replacing P_2 with P_3 in the expressions above. Adding the two together gives us the expectation numbers $E(j)$ of the five types of clusters seeded on day $\#i$. The sum $E = E(1) + E(2) + E(3) + E(4) + E(5)$ gives the total expectation number of clusters seeded on this day. In Ref. [6] we straightaway rounded this off to get the actual (integer) number of clusters seeded. Here however, there is one more thing to take care of, which is the maximum contact tracing capacity CT_{\max} .

Identification and isolation of a cluster entails tracking down 24 people and issuing quarantine recommendations to all of them. This is time- and resource-consuming work, and needs dedicated personnel. It is reasonable to expect that the city will have an upper limit on the number of contact tracers and hence a ceiling on the number of people who can be tracked down in a day. We call this ceiling CT_{\max} . Now we need an estimate of the number of people required to be traced every day. Generating a cluster of Types 2 through 4 requires identifying 24 people; the count required to generate a cluster of Type 1 is harder to calculate since it involves identifying UCT and SEC transmissions rather than cluster isolation. Nevertheless, for calculational convenience, we can assume approximately that this process involves tracing 24 people also. Generating a Type 5 cluster of course does not require any contact tracing at all. Putting this together, the number of people required to be traced on day $\#i$ is

$$n_{CT} = 24(E(1) + E(2) + E(3) + E(4)) = 24(E - E(5)) \quad . \quad (2)$$

If $n_{CT} \leq CT_{\max}$, then the calculated $E(j)$'s for day $\#i$ are feasible to generate. If this inequality is violated however, then the calculated $E(j)$'s are impossible.

In this case, we define the **capacity ratio** $\rho = CT_{\max}/n_{CT}$ and rescale the expectation numbers of contact traced clusters by ρ i.e. we replace $E(j) = \rho E(j)$ for j goes from 1 through 4. This rescaling ensures that the total number of contact tracings made on day $\#i$ becomes equal to CT_{\max} instead of exceeding it, while the relative proportions of Types 1 through 4 clusters remain unchanged. To make the rescaling consistent, we define $\rho = 1$ in the case that $n_{CT} \leq CT_{\max}$. The total number E of clusters seeded on day $\#i$ cannot depend on the contact tracers' ability to find these clusters, since the finding happens only after day $\#i$. Hence E itself is independent of ρ and the adjusted $E(5)$ is calculated as E less the rescaled $E(1)$ through $E(4)$. Only now do we perform the roundoff to calculate the integer numbers $\Delta z_i^{(j)}$ of susceptible clusters of type j seeded on day $\#i$.

Another difference from Ref. [6] occurs in the removal of immune clusters. In Ref. [6] we treated entire clusters as susceptible or immune – this was essential to prevent over- or undercounting while ensuring tractability of the mathematical expressions. Here, different types of clusters generate different numbers of

susceptible and immune people at the end of the cluster-level outbreak (Type 1 has 23 susceptible and 1 immune, Type 2 has 20 susceptible and 4 immune etc), and treating such clusters as a whole as immune (or susceptible) will lead to unacceptable levels of error. To get around this, we use an approximation scheme where the error involved is less. We let w_i and Δw_i (equivalent of z_i and Δz_i of Ref. [6]) denote the effective numbers of immune clusters. Whenever a Type 5 cluster is generated, we increase Δw_i by one, just as in Ref. [6]. For the other cluster types, whenever total 24 additional immune people are generated, we increase Δw_i by one. This scheme ensures that the infection can mathematically spread to the entire population, neither stopping short nor overshooting. Here, we explain why we opted for [1; 3; 6; 8; 5; 1] as the cluster sequence instead of [1; 3; 6; 7; 5; 1] of Ref. [6]. This is because, with this sequence, a Type 5 cluster infects everybody inside and is consistent with our heuristic susceptible cluster calculation scheme, whereas the one remaining susceptible person with the original sequence causes a needless headache at this step.

A final point of difference from Ref. [6] occurs in the treatment of the household cases. In Ref. [6], all socially active cases generate two additional household cases. Here, all at large cases will generate the two additional household cases but quarantined cases will not spawn this extra caseload.

§5. Computer algorithm. We now list the additional variables and parameters present in the model with respect to Ref. [6], and give the algorithm for calculating the case trajectories. In the algorithm, we condense those parts which have already appeared in Algorithm 1 of Ref. [6].

Variable	Significance
$x_i, \Delta x_i$	Cumulative respectively daily counts of at large cases cropping up on day $\#i$
$q_i, \Delta q_i$	Cumulative respectively daily counts of quarantined cases cropping up on day $\#i$
$w_i, \Delta w_i$	Cumulative respectively daily counts of effectively immunized clusters cropping up on day $\#i$. Replaces z_i and Δz_i of Ref. [6].

Parameter	Significance	Default value
P_0	Probability of random case being symptomatic	0.5
P_1	Probability of grounding a cluster upon finding a symptomatic case	0.8
P_2	Probability of identifying a symptomatic case's UCT secondary transmissions	0
P_3	Probability of identifying a symptomatic case's SEC transmissions	0.2
CT_{\max}	The maximum number of people whom the contact tracers can track down every day	50

Table 3 : New variables and parameters arising when contact tracing is added to the basic CST model.

We have chosen a relatively high default value for P_1 since clusters are composed of close contacts and these are relatively easy to trace. We have gone with zero as the default for P_2 since UCT events like transmission on board buses and inside marketplaces are almost impossible to identify in practice. For P_3 we have gone with a low default value since SEC transmissions, for example at wedding and birthday parties are harder to identify than intra-cluster transmissions but easier than transmissions in random public places.

We now give the algorithm itself. The algorithm is schematic and is not cast in any particular computer language; our code written in Matlab is freely available as discussed in the Declarations Section at the end of this Article.

Subroutine roundoff	
Carry over from Ref. [6]	
Main Routine	
Starting steps	
Step 1	Set parameters, initialize all variables with zero values
Step 2	for j goes from 1 to 5 Calculate $P_U(j)$ Calculate $P_S(j)$ end
Step 3	Set initial conditions to seed the system
Primary loop over i	
Step 1	Define $\alpha = \Delta x_{i-1} + \Delta x_{i-2} + \Delta x_{i-3}$ Break loop if $\alpha = 0$ beyond the seeding phase
Step 2	Calculate E_U as in Ref. [6], use w in place of z
Step 3	for j goes from 1 to 5 Calculate $E_U(j) = P_U(j)E_U$ end
Step 4	Calculate E_S as in Ref. [6], use w in place of z
Step 5	for j goes from 1 to 5 Calculate $E_S(j) = P_S(j)E_S(j)$ end
Step 6	for j goes from 1 to 5 Calculate $E(j) = E_U(j) + E_S(j)$ end Define $E = E(1) + E(2) + E(3) + E(4) + E(5)$
Step 7	Calculate $n_{CT} = 24(E - E(5))$
Step 8	if $n_{CT} > CT_{max}$ Define $\rho = CT_{max}/n_{CT}$ else Define $\rho = 1$ end for j goes from 1 to 4 Set $E(j) = \rho E(j)$ end Set $E(5) = E - E(1) - E(2) - E(3) - E(4)$
Step 9	for j goes from 1 to 5 Define $\Delta z_i^{(j)} = \text{roundoff}(E(j))$ end
Step 10	Set $\Delta x_{i+5} = \Delta x_{i+5} + (\Delta z_i^{(2)} + \Delta z_i^{(3)} + \Delta z_i^{(4)} + \Delta z_i^{(5)})v(1)$ $\Delta x_{i+10} = \Delta x_{i+10} + (\Delta z_i^{(3)} + \Delta z_i^{(4)} + \Delta z_i^{(5)})v(2)$ $\Delta x_{i+15} = \Delta x_{i+15} + (\Delta z_i^{(4)} + \Delta z_i^{(5)})v(3)$ $\Delta x_{i+20} = \Delta x_{i+20} + \Delta z_i^{(5)}v(4)$ $\Delta x_{i+25} = \Delta x_{i+25} + \Delta z_i^{(5)}v(5)$ $\Delta x_{i+30} = \Delta x_{i+30} + \Delta z_i^{(5)}v(6)$
Step 11	Set $\Delta q_{i+5} = \Delta q_{i+5} + \Delta z_i^{(1)}v(1)$ $\Delta q_{i+10} = \Delta q_{i+10} + \Delta z_i^{(2)}v(2)$

contd.

	$\Delta q_{i+15} = \Delta q_{i+15} + \Delta z_i^{(3)} v(3)$
	$\Delta q_{i+20} = \Delta q_{i+20} + \Delta z_i^{(4)} v(4)$
Step 12	Define $\Delta y_i = \Delta x_i + \Delta q_i$
Step 13	Define
	$imm = (\Delta z_i^{(1)} + 4\Delta z_i^{(2)} + 10\Delta z_i^{(3)} + 18\Delta z_i^{(4)})$
	$s_1 = imm \bmod 24$
	$\Delta w_i = \Delta z_i^{(5)} + (imm - s_1)/24$
	$s_2 = s_2 + s_1$
	if $s_2 \geq 24$
	$s_2 = s_2 - 24$
	$\Delta w_i = \Delta w_i + 1$
	end
Step 14	Set
	$x_{i+1} = x_i + \Delta x_i$
	$q_{i+1} = q_i + \Delta q_i$
	$y_{i+1} = y_i + \Delta y_i$
	$w_{i+1} = w_i + \Delta w_i$
Final steps	
Step 1	Define
	$\Delta f_{i+5} = 2\Delta x_i$
	$f_{i+5} = 2x_i$
Step 2	Prepare relevant plots

Algorithm 1 : Schematic representation of the routine used to compute case trajectories. Note that v denotes the cluster vector, which is $[1; 3; 6; 8; 5; 1]$ in many of the runs used here. imm in Step 13 of the primary loop counts immune population. The numbers 1, 4, 10 and 18 appearing in Step 13 might be replaced by $v(1)$, $v(1)+v(2)$, etc for greater generality.

As in Ref. [6], we introduce the parameter k_{\max} for convenience and set its value to 80. We use the initial condition that eight clusters are seeded on the first day, after which the disease evolves on its own.

---- 0 ----

RESULTS

§6. Complex effects in the model. In Ref. [6] we have already shown that the basic CST model has a plateau (i.e. extended period of nearly constant daily case rate) as a generic solution. We have also shown the **cryptogenic instability** where an increase in n_s causes the epidemic to increase very rapidly in severity, even though the contribution of the SEC events to the population-averaged contact rate remains quite small. In Fig. 3 we present a more detailed characterization of this instability beyond what was done in Ref. [6]. Considering the default values of Table 1 and no contact tracing, we vary n_s and plot two things : (a) the cumulative caseload at the end of the outbreak, and (b) the total duration of the outbreak. We show the caseload as a blue line associated with the left hand y -axis and the duration as a green line associated with the right hand y -axis. The actual points where we have run Algorithm 1 are shown as squares; the continuous lines are generated from them via interpolation using Matlab's "makima" routine. These conventions shall continue to hold for all Figures where we perform parameter sweeps.

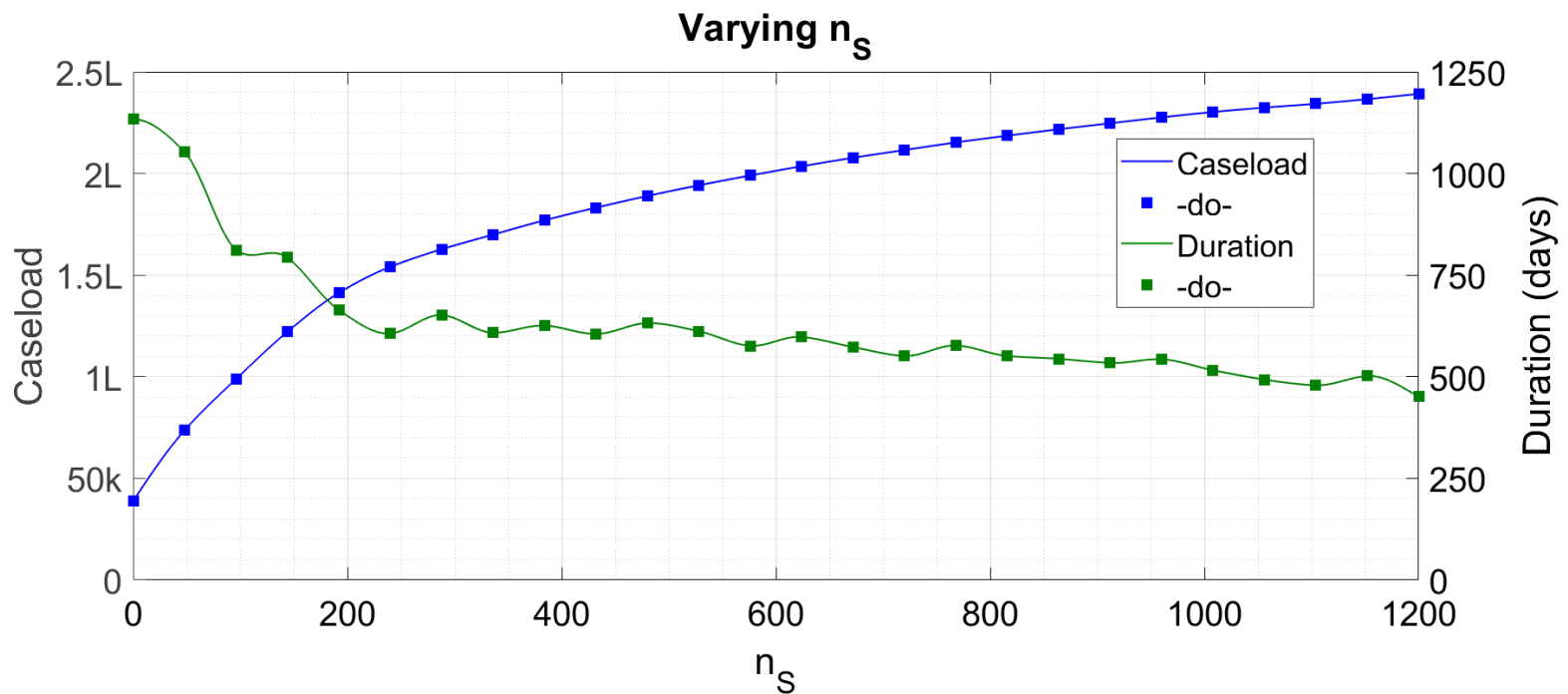


Figure 3 : Cumulative caseload and duration as n_s is varied with all other parameters set to default values. The symbol 'k' denotes thousand and 'L' lakh or hundred thousand.

We can see that the caseload initially increases rapidly with increase in n_s , and thereafter the rate of increase becomes lower.

In Ref. [6] we had identified another non-classical effect called **critical mass effect**. This refers to the fact that even when the parameter values are chosen to generate a wave, a sufficiently small initial condition results in the epidemic terminating quickly and at very few cases. We had demonstrated this effect in Fig. 4 of Ref. [6]. With the contact tracing included, we find that this effect is reinforced. For example, using the default values of Tables 1 and 2 and the value $n_s = 1800$, we get the following epidemic curve, Fig. 4. We show the daily case rate as grey bars.

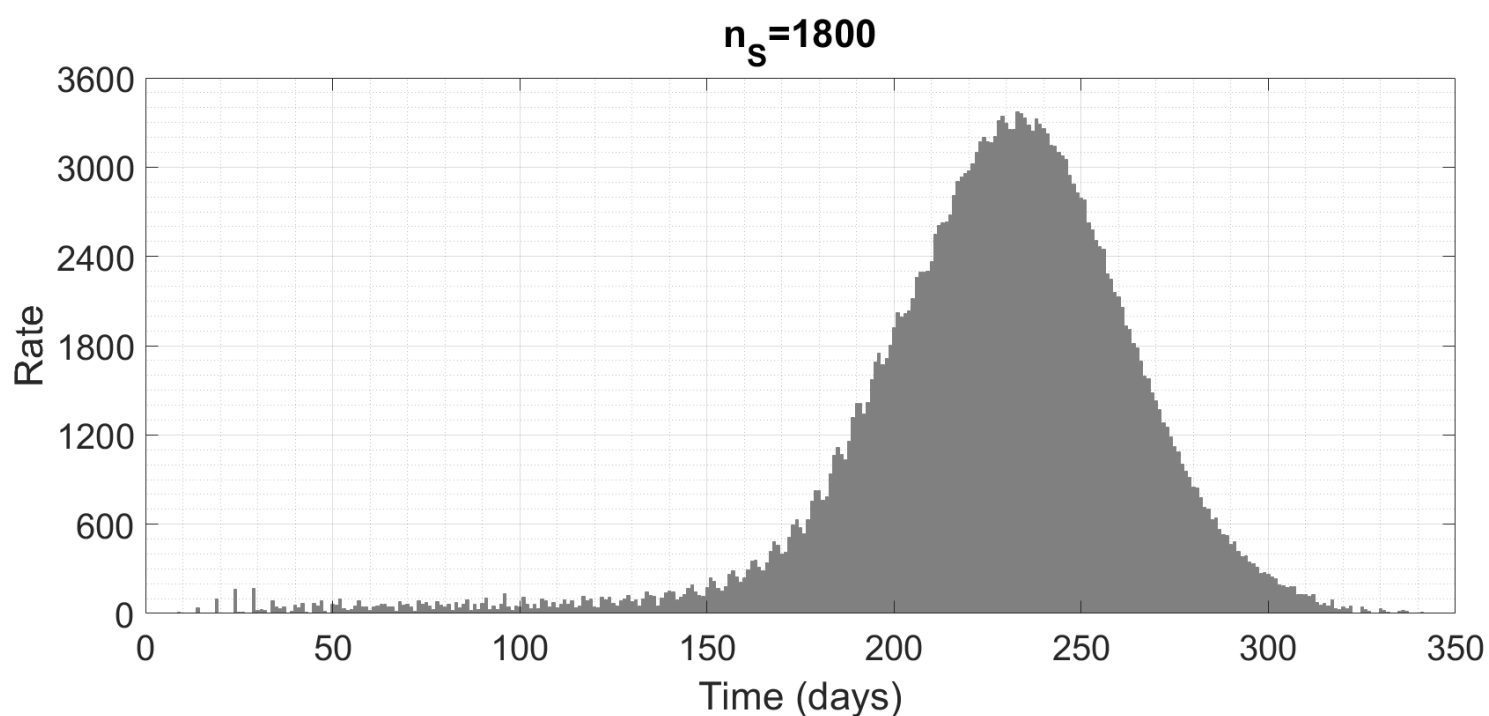


Figure 4 : Time trace of the epidemic showing the critical mass effect.

We highlight that this *entire* trajectory – with a nearly constant case rate from 40 to 130 days and a gigantic wave after 150 days – has been generated with parameters kept invariant throughout. Thus, the parameter values are clearly chosen to result in uncontrolled disease transmission; even so however, the epidemic continues at a simmering level for a very long time before the explosion occurs. This is the enhancement of the critical mass effect which takes place when contact tracing is added to the basic CST model. The

effect is no longer dependent on the initial condition's being sufficiently small but occurs much more generally.

For the next effect, we perform some parameter sweeps. Taking the default values from Tables 1 and 3 for all but one parameter, we vary this last remaining parameter from a low to a high value and plot the cumulative caseload and duration as in Fig. 3.

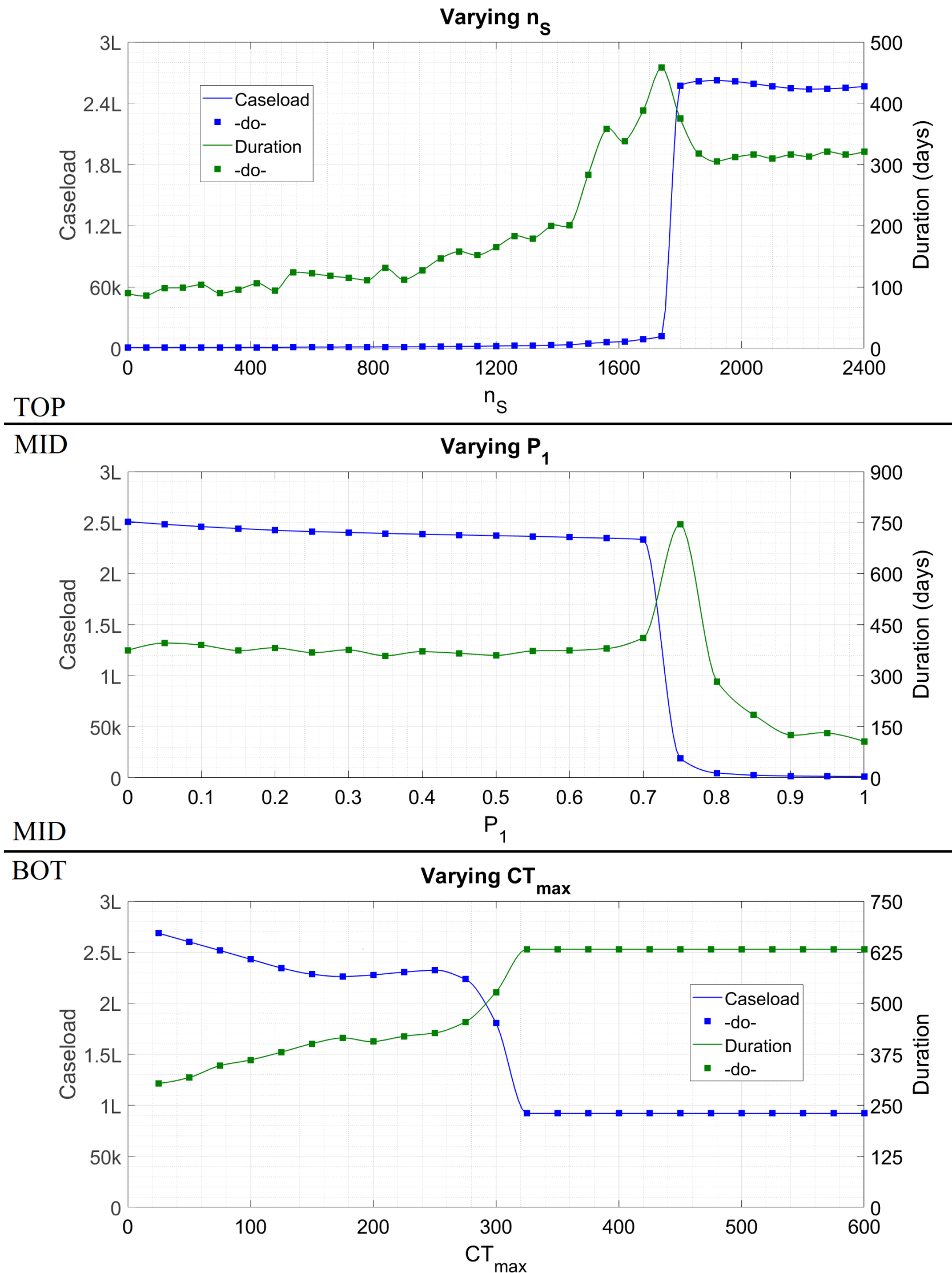


Figure 5 : Cumulative caseload (blue) and duration (green) of the epidemic as the labelled parameter is varied and all other parameters held to their default values. The symbol 'k' denotes thousand and 'L' lakh or hundred thousand.

In each plot we can see a step-function behaviour especially in caseload – at first it changes very gradually as the parameter is increased before abruptly showing a very sharp rise or drop at some critical value of the parameter. After this sudden transition, the variation is again gradual. On the happy side of this transition,

the time trace of the epidemic is a plateau while on the unhappy side it is a wave. Note that, on this wrong side of the transition, the epidemic proceeds to herd immunity. Thus, a knife-edge separates the tractable solutions from the intractable ones, and we call this phenomenon the **knife-edge effect**.

For the last effect, we consider time-varying parameters. We can see in Fig. 5-top that, with all other parameters set to default values, $n_s = 1900$ leads to an explosion while $n_s = 1500$ leads to controlled outbreak. Accordingly, a minimally invasive mitigation measure might involve reducing n_s from 1900 to 1500 after the outbreak is initiated. The results of this move, with the reduction implemented at day #100, are shown in Fig. 6-top; the results of a similar but more drastic reduction – n_s slashed to 300 instead of 1500 – are shown in Fig. 6-bot. In these and similar Figures, we show the daily case count as grey bars attaching to the left-hand y -axis and the cumulative caseload as blue line associated with the right hand y -axis.

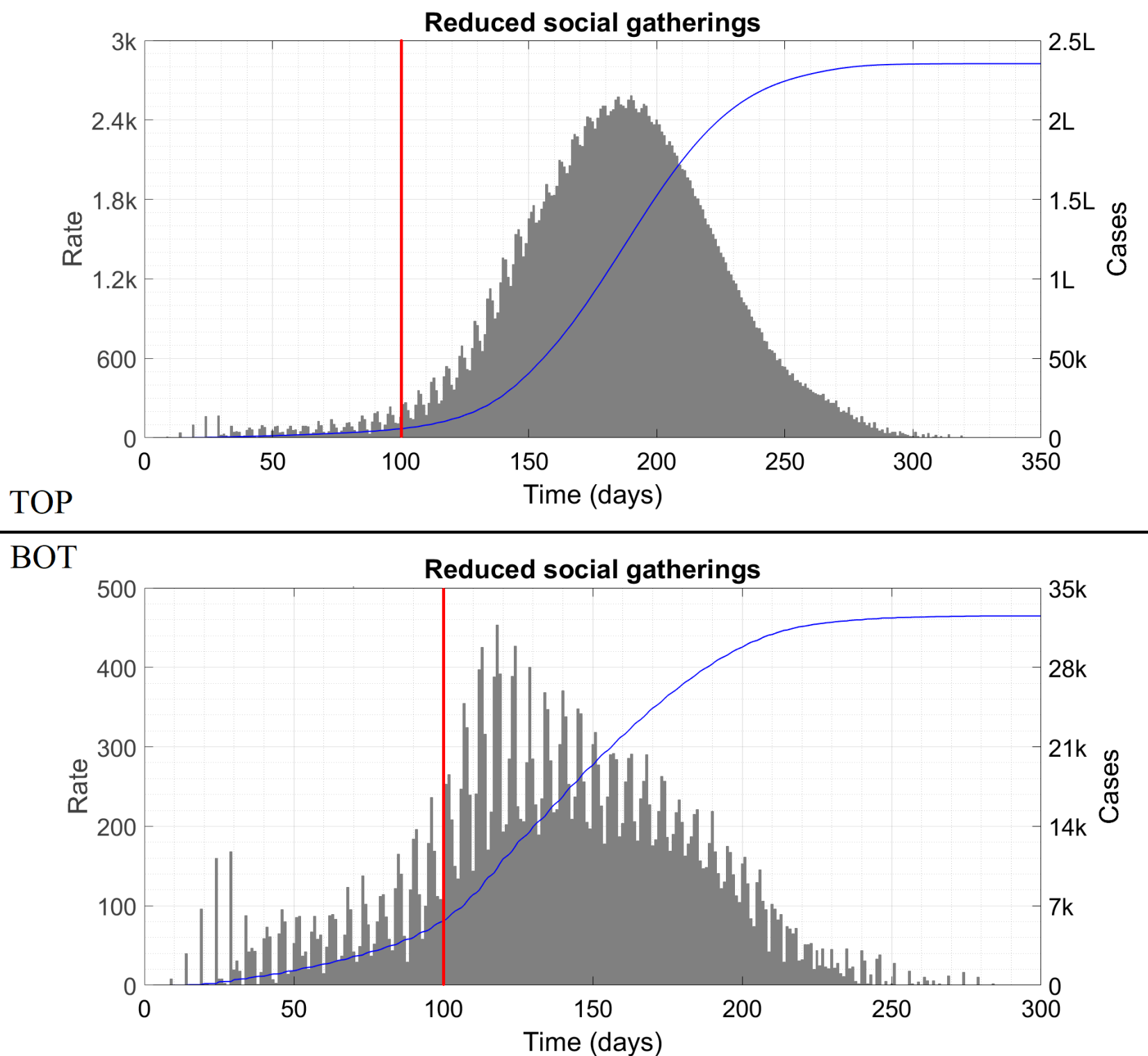


Figure 6 : Top panel shows the effects of reducing n_s from 1900 to 1500 while bottom panel shows the effects of reducing it to 300. In both cases, the intervention is applied at 100 days, indicated by the red line. The symbol 'k' denotes thousand and 'L' lakh or hundred thousand.

We can see that the weak intervention, although strong enough to prevent a serious outbreak if applied at the outset, has precisely zero effect when applied at a later stage. Acting later, a much stronger mitigation measure is necessary to bring the spiralling outbreak back under control.

To take another example, we start off as in Fig. 6, and consider the intervention of changing CT_{\max} rather than n_s . We implement a step increase from the value 50 to the value 400. The two panels of the below Figure show this intervention being applied at 120, respectively 100 days.

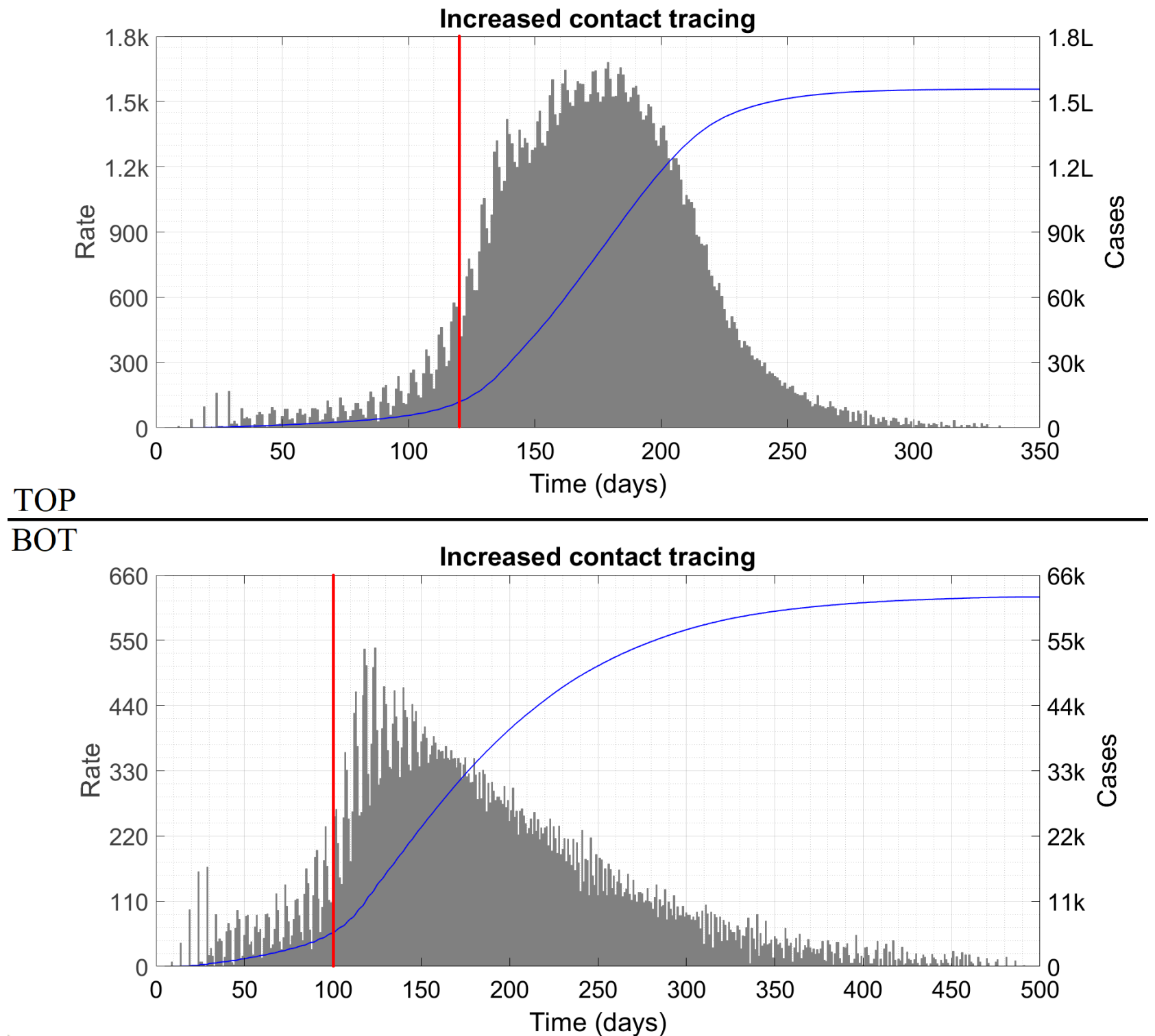


Figure 7 : Effect of increasing CT_{\max} from 50 to 400 at day #120 (top panel) and #100 (bottom panel). In both cases the intervention is depicted as a red line. 'k' denotes thousand and 'L' lakh or hundred thousand.

When applied at the 120-day mark, the intervention has little effect – the cases climb rapidly even after the measure is applied, peaking approximately 60 days after the change in policy. The effect of the measure is not zero since the cumulative count of approximately 1.5 lakh cases is less than the 2.5 lakh which would have accumulated in the absence of the intervention. There is a dramatic change however when the timing of the intervention is brought forward by 20 days. This time the cases peak about 20 days after the intervention, which is consistent with the time it takes for the infection to sweep through the clusters already seeded prior to the intervention. Thereafter however, the cases decrease slowly until the outbreak ends.

In both these Figures, we can see that the trajectory of the epidemic following an intervention depends not only on the strength of the intervention but also on the trajectory prior to the intervention. In a physical system, this phenomenon where the past state of the system influences its present behaviour is called hysteresis. The most famous example is magnetic hysteresis while similar behaviour in the forced Duffing

oscillator and other nonlinear systems is also pretty well-known. In the context of our epidemic model, we call this the **hysteresis effect**.

---- O ----

DISCUSSION

§7. Discussion of the various effects. We take the effects in the order presented. We have already spent considerable time on the cryptogenic instability in Refs. [6,7] so our treatment here will be brief. In Fig. 3, at the point $n_s = 1000$, the contribution of the SEC events to the population-averaged interaction rate is only 1 percent of the contribution of intra-cluster interactions; yet, the caseload in this situation is increased by six times. The average contact rate is a very important parameter in lumped parameter or compartmental models and a 1 percent increment in its value cannot cause such a huge increase in cases.

Critical mass effect in the CST model without contact tracing has also been identified in Refs. [6,7]. The tracing however is responsible for enhancing it substantially. In the absence of tracing, for the same parameter values we can get either an isolated handful of cases (for a small initial condition) or a full-blown wave (for a larger initial condition) but not both in series. With the tracing included however, Fig. 4 shows just such a solution and, as already mentioned in §1, it is a hallmark of COVID-19 trajectories round the world. We can also see that critical mass effect can result in a wave occurring months after a relaxation, when nobody is expecting trouble, as happened in India, Taiwan and elsewhere.

The qualitative reasoning behind the solution of Fig. 4 is as follows. Firstly, we note that contact tracing dramatically reduces the number of at large cases. With P_0 to P_3 at their Table 3 defaults, excluding the chance that a cluster is caught at seeding itself, there is a 24 percent probability that the cluster develops just one at large case and another 21 percent probability that it develops four at large cases, instead of 24. Now, in Fig. 4, the parameter values happen to be chosen in such a way as to generate a plateau if the contact tracers can access every emerging symptomatic case (note that the ‘if’ condition is not actually satisfied by the situation at hand). After the seeding period (first 40 days), the number of fresh cases occurring per day is such as to stretch the tracing capacity to its limit or very slightly beyond. Thus, for a long time, nearly all feasible cases are traced and the plateau continues with a very slight increasing trend (visible when zoomed up, which will come later). The plateau is not perfectly smooth but consists of fluctuations about a constant envelope. The increasing trend means that, as time goes on, days of locally high cases definitely seed more clusters than the tracers can access and isolate. These untraced clusters in turn start supplying considerable numbers of at large cases. Now however we enter a vicious circle, for more at large cases with constant contact tracing capacity means yet more untraced clusters and at large cases. This vicious circle manifests as the sudden wave. Thus we can say that the emergence of the wave is the result of a feedback process between the contact tracing and the case trajectory itself.

Qualitatively, this destabilization of the plateau by itself is somewhat similar to the phenomenon of autoparametric resonance in mechanical systems [8] where the motion itself activates the instability which destroys it. We have observed that increasing n_s above the value used in Fig. 4 results in the duration of the preliminary plateau phase reducing rapidly. This is consistent with our explanation since higher n_s saturates the contact tracing system and leads to the vicious circle faster. A mathematically rigorous analysis of the critical mass effect in the epidemic model is left by us for future study. For now, the important thing is that the critical mass effect has played a key role in affecting the COVID-19 epidemic trajectories worldwide. It is at best a thorn in the side and at worst a nemesis of public health planners, since, right until the wave occurs, there is scant evidence to suggest that it is imminent and that an increase in NPI or compliance therewith is urgently called for.

Like the critical mass effect, knife-edge effect confounds the efforts of epidemic predictors and planners. The risk here is that the transition from tractable (plateau) to intractable (wave) behaviour is nearly instantaneous, and the qualitative behaviour on either side of the knife-edge is the same. Hence, if a region is operating in plateau mode with a certain level of NPI, it is impossible to guess how much further NPI

relaxation is possible without the plateau being destabilized. Incremental relaxation – a policy used almost worldwide during the initial exit from lockdown – will scarcely produce results until the knife-edge is actually hit, when the consequences will become drastic. This effect is similar to phase transitions in fundamental physical as well as in more applied processes [9-12], of which our model can be considered a certain type. Once again, the mathematical intricacies and parallels with other branches of science must be allocated the loop line for now, with preference going to the ramifications of the effect on COVID-19 explanation and prediction. We believe that the unexpected undulations in caseload, in for example Fig. 5-top for $n_s > 1800$, are the result of a numerical error and not significant.

The hysteresis effect is almost self-explanatory, so we proceed straight to a real-world example. The state of Maharashtra in India, including its two primary case-driver cities Mumbai and Pune, is an excellent demonstrator of this effect. During January and February 2021, cases in these cities were constant or decreasing even though NPI were not very stringent (there weren't too many restrictions in place beyond a mask mandate and size cap on gatherings, and even these weren't being followed too strictly). When the wave arose, the reproduction number R was relatively low [7] and the State Government initially tried a soft intervention approach such as imposing night curfews and increasing mask compliance. As per compartmental epidemic models, this alone should have worked – if R was below unity in January at a certain level of NPI, then reverting to that level of NPI in March should have brought it below unity again and stemmed the growth of cases. This however was not what happened – the daily case counts grew relentlessly and only when a full lockdown was imposed did they start going down again. Hysteresis effect can explain this observation completely.

All of these effects combine together to explain the bewildering variety of COVID-19 trajectories seen all over the world. Anomalously high case counts, for example in USA and Israel, might indicate the crossing of a knife-edge during the relaxation process or an ineffective contact tracing system while anomalously low counts, for example in India and UK, might indicate a better contact tracing system and operation on the bright side of a knife-edge. (A recent claim [13] of India's having reached herd immunity with 90 crore infections as of 30 June 2021 is based on imaginary and not real data, and a model named 'SEIR-fansy'; this nomenclature is singularly appropriate.) The complex effects underlying corona trajectories might also influence analyses of transmissibility of mutants, especially when this analysis is done with respect to a compartmental model. For example, the transmissibility of B1.1.7 variant appears to have been overestimated since it was eliminated in competition while wildtype survived. While B1.617.2 is definitely more transmissible than other variants, some accounts of its infectiousness (for example, it is like the chickenpox [14], it has an R_0 upto 8 [15]) might be exaggerated (recall that not every country where this variant is prevalent is reeling under its attack). This will be especially true is the serial interval for this variant turns out to be lower than for the other variants, which one study [16] suggests to be the case. The epidemic doubling time T_d during exponential growth is defined as $T_d = (\log 2) T_s / \log R$ [5] where T_s is the serial interval; a smaller T_s results in faster doubling at the same R . The question of whether the transmissibility of B1.617.2 is overhyped due to considerations other than scientific ones cannot be answered and is best not asked.

§8. Model limitations and their consequences. Most of the limitations of the model have already been mentioned in Ref. [6]. They are

- Assumption of constant cluster size
- Assumption of constant cluster sequence
- Use of roundoff
- Decouplement of households from clusters
- Ignoring the possibility of one person belonging to multiple clusters

The ways of circumventing these assumptions have also been discussed in Ref. [6] – in short, we have to convert the deterministic model to a fully agent-based model. The additional part implementing contact tracing contains two assumptions which we have already discussed in §4. There is an approximation in the calculation of ρ since we assume that all the cluster members are being traced on the same day as the case

is found. In reality, the most immediate contacts might be identified first and the more distant ones later, easing the tracers' job. On the average however, the total number of contacts needing to be traced over say a week or ten days will remain unchanged. Another approximation arises in our counting of immune clusters w_i and Δw_i . In this case, since our scheme preserves the total susceptible and immune populations at any time, errors will tend to average out.

Once again, converting from the deterministic CST to a fully agent-based model will obviate the need for most or all of these assumptions. A more interesting question is : how might our assumptions influence the model results ? As a first test, we have varied the parameter which might be the most restrictive – the cluster sequence. Instead of [1; 3; 6; 8; 5; 1], we have now considered the sequence [1; 4; 13; 6]. This describes much faster intra-cluster transmission, corresponding to a more transmissible mutant. Concomitantly we have also increased the values of P_U from 0.15 to 0.25 and m_s from 2 to 3.5. Note that these values are hypothetical and we neither state nor intend to suggest that the percentage increases in transmissibility parameters are representative of B1.617.2 double mutant. In this case, we have run extensive simulations and found *all* the effects mentioned in §6. The parameter values at which the different behaviours occur are of course different, but that apart everything else is the same. We dispense with another plethora of Figures, and instead make the code available to whosoever wants it (see Declarations).

A realistic society is expected to have a distribution of cluster sizes and sequences as well as transmission parameters, and not just constant values of all of them. At one end of the spectrum, college students will tend to have large clusters, high intra-cluster interaction resulting in a rapid sequence, as well as higher values of P_U and m_s arising from their active social lives (and possibly lack of caution). At the other end, retirees are expected to have small clusters, a less sharp cluster sequence and low P_U and m_s due to their relatively limited social lives as well as justified fears about the consequences of contracting the disease. In such a society, we expect some of the special effects to remain more or less as is and others to be altered.

The critical mass effect ought to be robust because of its mechanism of action. With parameters in a dangerous region, when the college students are operating at the borderline of tracing capacity, cases in that sub-population will plateau; cases in less contagious sub-populations will also plateau or decay (recall that in CST model, the plateau is manifest in large regions of the parameter space). The students will be the first to cross the contact tracing threshold and activate the instability; cases arising there will then induce cases in the other sub-populations and cause them to follow suit as well. On the other hand, a distribution of cluster size and contagiousness will certainly blunt the edge of the knife. When an NPI is progressively relaxed, the critical parameter value for the college students will be lower than that for the retirees, and crossing that first cutoff alone will not initiate a wave in the whole population. For a mass transition from plateau to wave, we shall have to effect a larger change in parameters than Fig. 5 would imply. Like critical mass effect, hysteresis effect is again expected to be robust in a population with a distribution of cluster sizes and parameters, since it will apply individually to the various sub-populations. As we have mentioned in §7, demonstrations of the critical mass and hysteresis effects have indeed occurred many times in different parts of the globe. Finally, we mention that the consequences of errors arising from approximations in our implementation of the model (for instance, using the parameter k_{\max}) will be no different from those of approximations in the model itself.

§9. Early warning indicator. The complicated effects we have described in §6-7 are certainly capable of explaining the bewildering epidemic trajectories we are seeing all over the world. They do however raise the question, is everything upto chance ? Is there no way of opening up, assessing in real time whether there really is elevated risk of a wave and then rolling back if necessary before it is too late ? To prevent a wave, is our only option to follow a path of maximum caution although by now even the most law-abiding citizens are sick of COVID-appropriate behaviour ?

A path of less caution can be followed only if there are early-warning indicators of a forthcoming wave, which can signal the imminency of the wave during the quiescent phase. For an alarm to be reliable, it must 'go off' every time a wave is really impending, and preferably not cry wolf when there is no danger. Indicators derived from disease trajectories alone are inadequate since (as we have seen above) the trajectories depend on too many things at once and mask important details of the processes through which

they were generated. Indeed, a detailed discussion of this point has been presented by Dablander et. al. [17] who conclude that many potential early warning indicators actually decrease rather than increasing before a wave, and hence are useless.

In our analysis however, we have found one parameter which consistently appears to predict the onset of a wave. This is the capacity ratio ρ i.e. the ratio of the maximum contact tracing capacity of the city's healthcare personnel to the actual number of contacts who are being traced per day. We have found that so long as ρ is strictly unity, there is no question of a wave, while if ρ becomes less than unity, then there is a very real threat (recall that ρ is constrained to a maximum value of unity by definition). To appreciate the role of this parameter, we go back to our understanding (§7) of a wave as a vicious circle between emerging cases and untraced clusters. While $\rho = 1$, all feasible contacts are getting traced, i.e. the number of potential contacts emerging per day is less than CT_{\max} . This automatically puts a cap on the maximum number of daily cases. Hence, there is no question of a wave while $\rho = 1$. On the other hand, when ρ decreases below unity, there is a surplus of at large cases which kicks off the vicious circle. We expect that, during the initial phases at least, the surplus will be small – it will take some time for the feedback loop to amplify the untraced cases to a level where the growth becomes uncontained. During this time, the actual case counts will remain small but ρ will become less than unity, thus acting as an early warning indicator.

To verify our hypothesis, we now present extensive simulation results. As a first example, we consider Fig. 4. Here, the parameter value $n_s = 1800$ generated a huge wave after a faux plateau, while if we reduce n_s to 1750 then we get the plateau only (note from Fig. 5-top that the knife-edge for this situation lies between 1750 and 1800). In the below Figure, we show the epidemic history for the two situations respectively during the initial 120 days. We also show ρ as a blue line attaching to the right hand y -axis.

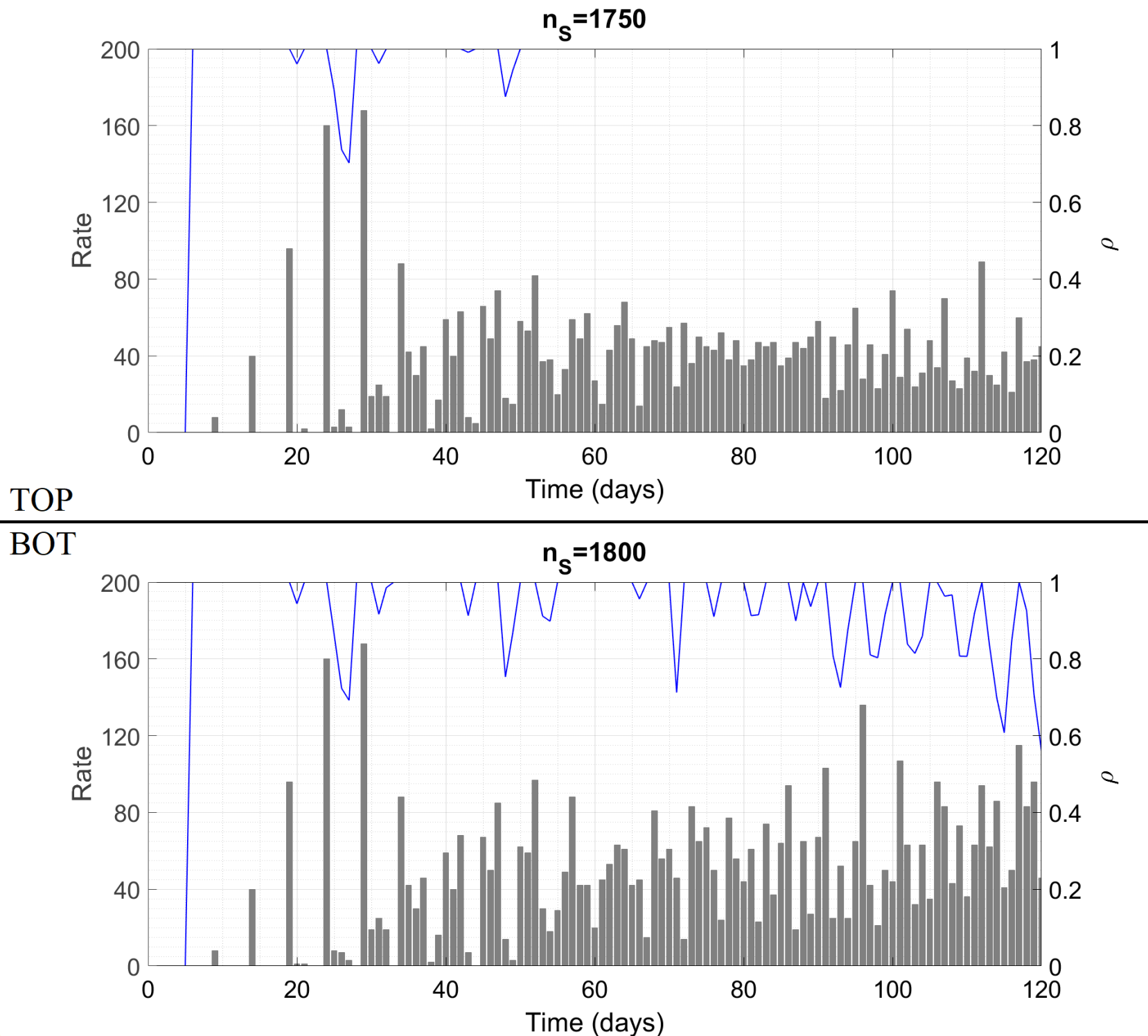


Figure 8 : The initial portions of case trajectories and capacity ratio as a function of time for two contrasting situations.

Indeed, the case trajectories in the two panels look nearly identical. But while ρ for the top panel is unity almost throughout (the dip at around 30 days is related to the initial condition rather than the steady state dynamics), ρ for the bottom panel shows marked deviations which increase as time goes on. The bottom panel does show a very slight increasing trend superposed on the plateau (see again §7) but this can easily be missed or taken for an insignificant fluctuation. The fact that the absolute case counts are higher in the bottom panel than in the top is insignificant since a plateau is a plateau, and absolute case counts have no meaning. But the trend in ρ in the two plots is very different, and this difference is very significant as the subsequent evolution demonstrates (from Fig. 4 we can see a gradual rise in cases between day #120 and #150, and a meteoric rise thereafter). Taking cognizance of the decrease in ρ around day #100 however, if the public health authorities take action, then the horrors after day #150 can be entirely avoided.

As a second example we consider the two situations of Fig. 7 which demonstrated the hysteresis effect. In both situations, the case counts initially increased after the intervention was applied; only in the adverse situation however was there a full-blown wave despite the intervention. Can the marker ρ again predict the difference ?

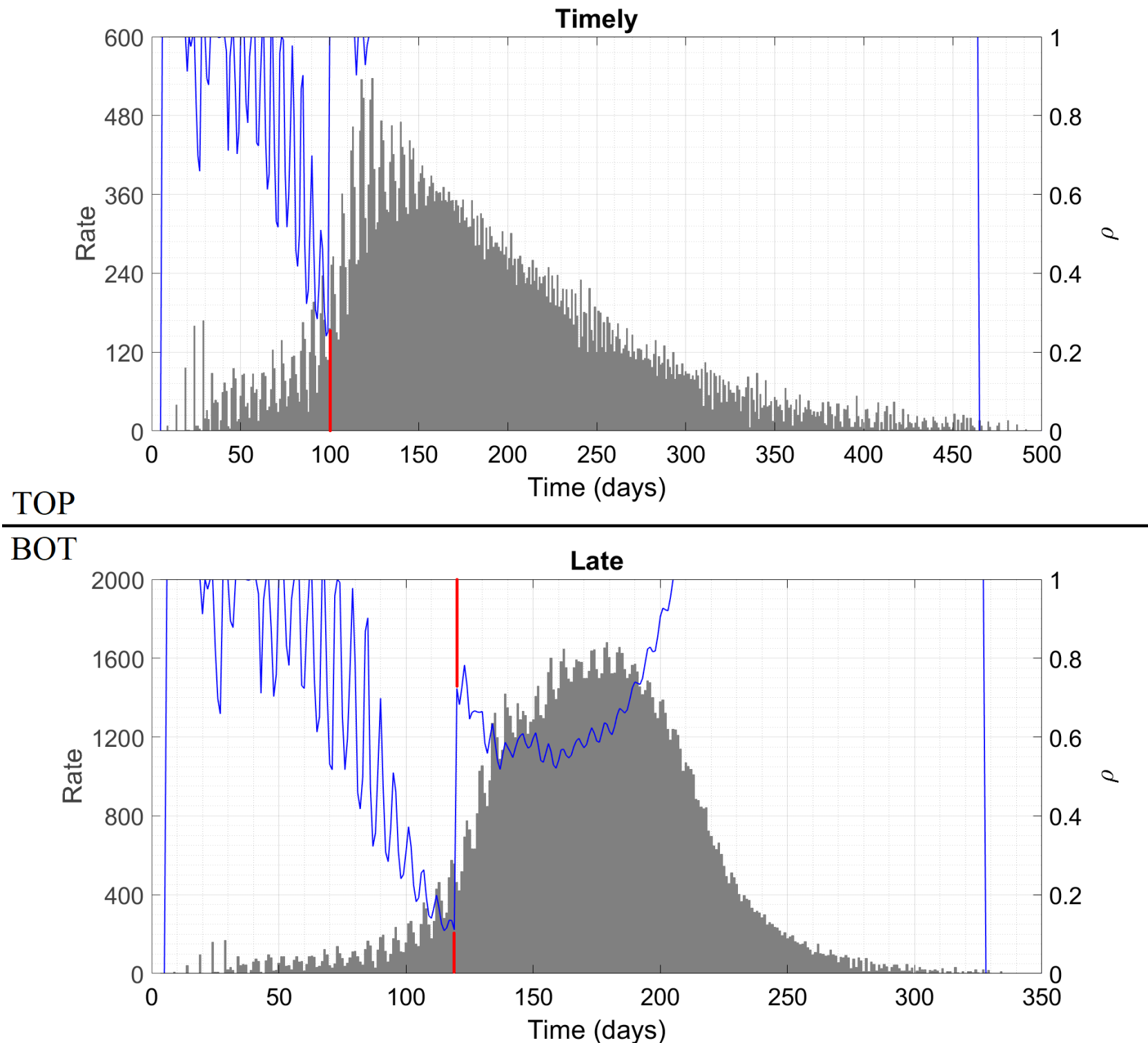


Figure 9 : Case trajectories and capacity ratio as a function of time for two contrasting situations. The red lines indicate the interventions, as in Fig. 7. We have shown only parts of the red lines so that they do not overlap with the blue line for ρ which has a jump exactly at the point of application of the intervention.

You bet. When the intervention is timely, ρ goes up to unity immediately. There are just two isolated departures and that is it. The tracers isolate enough cases to starve the epidemic of fuel and it proceeds slowly but inexorably towards its end. When the intervention is late however, ρ climbs only to about 70 percent, still leaving scope for a considerable number of untraced, at large cases. As the wave continues, ρ dips still further until the epidemic peaks on its own, ρ hits the ceiling and then the disease rapidly vanishes.

As a third example, we consider the rogue mutant which has the cluster sequence [1; 4; 13; 6] and the values $P_U=0.25$ and $m_S=3.5$. We also increase CT_{\max} to 100, both to change the parameter from Fig. 8 and also to incorporate the fact that a rogue mutant will need to be combated by a larger contact tracing force. Keeping P_0 to P_3 the same as before, when n_S is varied, the knife-edge occurs between 750 and 800. In the latter case, there is a very pronounced demonstration of the critical mass effect. Below we plot the initial 200 days of the case trajectories and ρ for these two values of n_S .

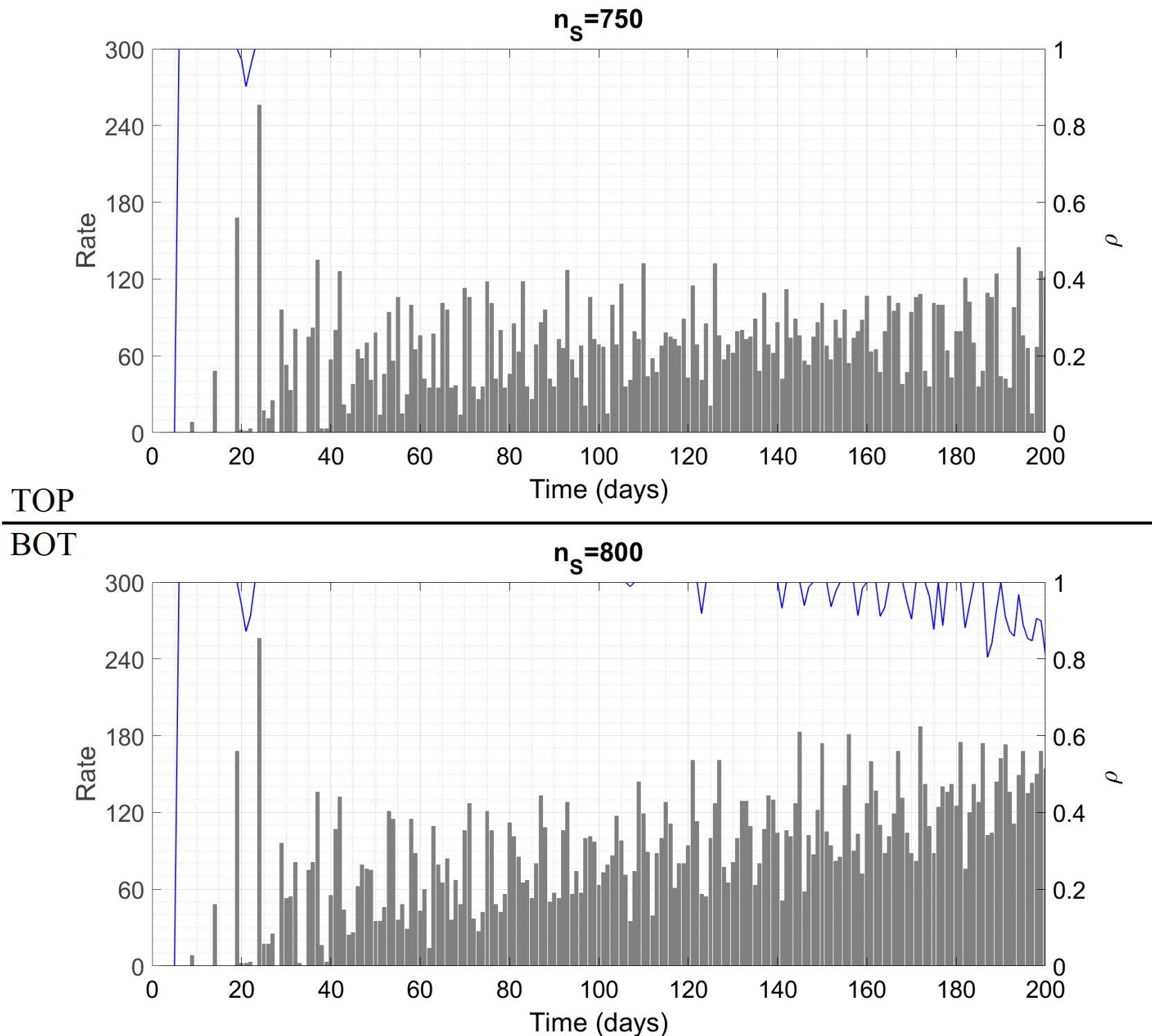


Figure 10 : The initial portions of case trajectories and capacity ratio as a function of time for two contrasting situations.

This is almost identical to what we saw in Fig. 8. In the subsequent evolution not shown in the Figure, the top panel proceeds almost unchanged to a cumulative caseload of 32,000 while the bottom panel takes off at around day #230 to reach a maximum daily rate of 2500 cases and a cumulative caseload of almost 2,30,000. A difference with Fig. 8 is that ρ for the adverse situation here is higher than that in Fig. 8-bot. Due to higher transmissibility of the mutant however, this smaller surplus of at large cases is sufficient to create a situation from which there is no recovery. Also, CT_{\max} itself is larger here, so a smaller percentage deviation from unity corresponds to a larger absolute count of untraced cases. For this reason, the only acceptable value for ρ is strictly unity and nothing less; this is not too stringent a condition since ρ cannot exceed 1 by definition.

These three sets of simulations, together with the logic underlying its function, provide convincing evidence that the capacity ratio ρ indeed has the power to act as an early warning indicator of an imminent wave. A less-than-unity value of ρ is practically difficult to measure since the denominator n_{CT} can only be estimated or calculated in this case (when $n_{CT} \leq CT_{\max}$, it can of course be measured). When $\rho < 1$, what can be measured as a substitute is the ratio of cases who are self-reporting symptoms to those being identified in contact tracing drives. If this ratio increases, then trouble is indicated. Similarly, the average number of contacts being traced for each freshly reporting case can also be measured. By now, its value during quiescent periods is known; a lower ρ will correspond to a decrease of this number and signal the need for

more proactive measures. Finally, ρ can also be connected to the test positivity rate; for that we shall require data about the testing schedules of suspected contacts, which we do not have currently. Hence we leave this connection for a future study.

§10. Public health implications. Public health policy decisions must be formulated on the basis of results which remain qualitatively unaffected by changes in parameter values or refinements in the model structure. The first such result is that SEC events activate the cryptogenic instability and can cause a high caseload. For this reason, size limits on such gatherings must be imposed and enforced. Since SEC events contribute a small fraction to population-averaged interactions, the social impact of such a move will be much less than that of say closure of restaurants and bars, where the transmission is primarily intra-cluster. Thus, public health messaging may focus on narrowing the scope of social interactions instead of reducing their frequency.

The second robust conclusion is that contact tracing is of great importance in combating the spread of the disease. By preferentially isolating suspected cases, it can save mass isolations i.e. lockdowns and similar measures. Faster tracing with higher probability and higher capacity can lead to enormous socioeconomic gain with negligible disease cost. To facilitate the contact tracers' job, people may be encouraged to keep diaries in which they make lists of every person with whom they have had close, unmasked interactions. Then, in case a person turns out positive, the entries for the last five or seven days in his/her diary may be used to access potential contacts quickly. These diaries might also be used to trace a person's SEC transmissions and nip the seeded clusters in the bud. Compared with electronic contact tracing methods, this method has the benefit that one's contacts are not logged unless one is actually suspected of exposure, and there is no question of the diaries being used for surveillance without one's knowledge. A disease mitigation tool which supplements contact tracing is random testing i.e. testing of asymptomatic persons. Here we have not accounted for this intervention explicitly, but its effects can be analysed easily and the policy implemented as per cost and complexity considerations.

An extensive contact tracing program might have an unpleasant corollary in that many people might be treated as potential cases and quarantined despite not actually having been exposed to the virus. In our example of §4, if Team is a Type 1 cluster then only Bravo, Charlie and Delta are actual cases while Echo through Xray are also grounded out of abundance of caution. Thus, there are more than six times as many quarantine orders issued as are necessary. For high-interaction persons like shopkeepers, such false alarms might be raised every few days. This will cause economic harm, undue mental stress as well as public lack of faith in the contact tracing system. To mitigate this, two options are available. The first is for asymptomatic suspected exposures to perform a daily self-test, the logistics of which as a quarantine substitute have been discussed in Ref. [18]. The second is to not ground the person entirely unless symptomatic but to ensure that over the next few days s/he follows 100 percent adherence to COVID-19 protocols at work and essential activities, and refrains from inessential activities. Data from Cornell University's extensive case surveillance network has not found even one instance of in-classroom viral transmission, suggesting that a 'soft' quarantine may be as effective at preventing transmission as a 'hard' one.

The third conclusion which appears to be robust is that saturation of the contact tracing infrastructure might indicate the imminent arrival of a wave. When such saturation is observed, a rollback of reopening measures should be announced immediately and maintained until the strain on the contact tracers is eased. Augmenting the contact tracing capacity by hiring professional part-time and full-time tracers is a longer-term solution which will enable reopening.

§11. Conclusion and future directions. In this Article we have added contact tracing to the cluster seeding and transmission (CST) model of COVID-19 developed in Refs. [6,7]. We have found that by adding this feature, the model results become significantly more realistic. We have identified four effects – cryptogenic instability, critical mass effect, knife-edge effect and hysteresis effect – which make epidemic trajectories unpredictable and contribute to the perplexing variety of COVID-19 trajectories seen all over the world.

Then, we have tried to extract order from amidst this chaos. For this we have identified conclusions which are robust across parameter values and hold true in very general situations. The role of SEC events in spreading disease and the role of contact tracing in mitigating it belong to this class. Finally, we have proposed the capacity ratio ρ as an early warning indicator of an imminent wave – if this does turn out to be effective in reality, then we might all find collective peace of mind from the presence of a reliable advance detection system.

There are at least two directions for future work. The first is theoretical and consists of a comprehensive mathematical analysis of the various effects we mentioned here. Such analysis is bound to reveal fascinating connections between epidemic spreading and other branches of science and engineering. The second direction is practical. Since many countries are now racing ahead with vaccination drives, our near-term future plan has to be the analysis of this intervention. It is noteworthy that vaccination appears to be producing different kinds of results in different countries; part of this may be due to the different efficacies of the various vaccines in use while part may be due to the effects we have described here. Although COVID-19 is currently nowhere near elimination, there might still be hope in the longer term, when vaccines get developed with more durable and broad-spectrum immune responses. The neutralization of COVID-19 as a global threat, if at all possible, will have to be brought about through a combination of vaccination and public health efforts, and mathematical modeling will be essential to achieve their most effective synergy.

--- O --- O --- O --- --- O --- O --- O ---

DECLARATIONS

We have NO conflict of interest to declare.

We have NOT been funded by any agency to conduct this study.

Should you wish to access our Matlab codes, either for verification or for performing scenario analysis, please contact SHAYAK at sb2344@cornell.edu or shayak.2015@iitkalumni.org regarding the same. We shall comply with bona fide demands for code as fast as possible.

REFERENCES

- [1] <https://www.nature.com/articles/d41586-021-02125-1>
- [2] <https://clingen.igib.res.in/covid19genomes/>
- [3] <https://covid.cdc.gov/covid-data-tracker/#variant-proportions>
- [4] NG Davies, S Abbott, RC Barnard et. al., “Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England,” **Science** 372 (6538), eabg3055 (2021)
- [5] B Shayak and MM Sharma, “A New approach to the dynamic modeling of an infectious disease,” **Mathematical Modelling of Natural Phenomena** 16, 33 (2021)
- [6] B Shayak and MM Sharma, “A Cluster-based model of COVID-19 transmission dynamics,” **MedRxiv Article** (2021). Available at <https://www.medrxiv.org/content/10.1101/2021.06.02.21258243v1>
- [7] B Shayak and MM Sharma, “The Origins of the second wave of COVID-19 in India – a model-based investigation,” **Socarxiv Article** (2021). Available at www.osf.io/preprints/socarxiv/ce9kx
- [8] F Verhulst, “Parametric and autoparametric resonance,” **Acta Applicanda Mathematica** 70, 231-264 (2002)
- [9] B Drossel and F Schwabl, “Self-organized critical forest fire model,” **Physical Review Letters** 69 (11), 1629-1632 (1992)
- [10] L Palmieri and HJ Jensen, “The Forest fire model : subtleties of criticality and scale invariance,” **Frontiers in Physics** 8, 257 (2020)
- [11] PJ Reynolds, HE Stanley and W Klein, “Percolation by position-space renormalization group with large cells,” **Journal of Physics A** 11 (8), L199-L208 (1978)
- [12] SV Buldirev, R Parshani, G Paul, HE Stanley and S Havlin, “Catastrophic cascade of failures in interdependent networks,” **Nature** 464 (7291), 1025-1028 (2010)
- [13] L Zimmermann, S Bhattacharya, S Purakayastha et. al., “SARS-CoV-2 infection fatality rates in India : systematic review, meta-analysis and model-based estimation,” **MedRxiv Article** (2021). Available at

<https://www.medrxiv.org/content/10.1101/2021.09.08.21263296v1>

[14] <https://www.webmd.com/lung/news/20210730/war-has-changed-cdc-says-delta-as-contagious-as-chicken-pox>

[15] Y Liu and J Rocklov, "The Reproduction number of the delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus," **The Journal of Travel Medicine**, Online (2021)

[16] S Ryu, D Kim, JS Lim, ST Ali and BJ Cowling, "Changes in the serial interval and transmission dynamics associated with the SARS-CoV-2 delta variant in South Korea," **MedRxiv Article** (2021). Available at

<https://www.medrxiv.org/content/10.1101/2021.08.18.21262166v1>

[17] F Dablander, JAP Heesterbeek, D Borsboom and JM Drake, "Overlapping time scales obscure early warning signals of the second COVID-19 wave," **MedRxiv Article** (2021). Available at

<https://www.medrxiv.org/content/10.1101/2021.07.27.21261226v3>

[18] L Feretti, C Wymant, A Nurtay et. al., "Modeling the effectiveness and social costs of daily lateral flow antigen tests versus quarantine in preventing onward transmission of COVID-19 from traced contacts," **MedRxiv Article** (2021). Available at <https://www.medrxiv.org/content/10.1101/2021.08.06.21261725v1>