Low case numbers enable long-term stable pandemic control without lockdowns

Sebastian Contreras[1,2], Jonas Dehning[1], Sebastian B. Mohr[1], Simon Bauer[1], F. Paul Spitzner[1], and Viola Priesemann[1,3*]

[1]Max Planck Institute for Dynamics and Self-Organization, Am Faßberg 17, 37077 Göttingen, Germany.
[2]Centre for Biotechnology and Bioengineering, Universidad de Chile, Beauchef 851, 8370456 Santiago, Chile.
[3]Department of Physics, University of Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany.
*Corresponding author. viola.priesemann@ds.mpg.de
All authors contributed equally.

Abstract

The traditional long-term solutions for epidemic control involve eradication or population immunity. Here, we analytically derive the existence of a third viable solution: a stable equilibrium at low case numbers, where test-trace-and-isolate policies partially compensate for local spreading events, and only moderate restrictions remain necessary. In this equilibrium, daily cases stabilize around ten new infections per million people or less. However, stability is endangered if restrictions are relaxed or case numbers grow too high. The latter destabilization marks a tipping point beyond which the spread self-accelerates. We show that a lockdown can reestablish control and that recurring lockdowns are not necessary given sustained, moderate contact reduction. We illustrate how this strategy profits from vaccination and helps mitigate variants of concern. This strategy reduces cumulative cases (and fatalities) 4x more than strategies that only avoid hospital collapse. In the long term, immunization, large-scale testing, and international coordination will further facilitate control.

Introduction

As SARS-CoV-2 is becoming endemic and knowledge about its spreading is accumulated, it becomes clear that neither global eradication nor population immunity will be achieved soon. Eradication is hindered by the worldwide prevalence and by asymptomatic spreading. Reaching population immunity without an effective vaccine or medication would take several years and cost countless deaths, especially among the elderly (*1, 2*). Moreover, evidence for long-term effects ("long COVID") is surfacing (*3–6*), advising against strategies aiming to progressively exposing people to the disease so that they acquire natural immunity. Hence, we need long-term, sustainable strategies to contain the spread of SARS-CoV-2. The common goal, especially in countries with an aging population, should be to minimize the number of infections and, thereby, allow reliable planning for individuals and the economy – while not constraining individuals' number of contacts too much (*7*). Intuitively, a regime with low case numbers would benefit not only public health and psychological well-being but would also profit the economy (*8, 9*).

However, control of SARS-CoV-2 is challenging. Many infections originate from asymptomatic or pre-symptomatic cases (*10*) or indirectly through aerosols (*11*), rendering mitigation measures difficult. Within test-trace-and-isolate (TTI) strategies, the contribution of purely symptom-driven testing is limited, but together with contact tracing, it can uncover asymptomatic chains of infections. The rising availability of effective vaccines against SARS-CoV-2 promises to relieve the social burden caused by non-pharmaceutical interventions (NPIs). However, it is unclear how fast the restrictions can be lifted without risking another wave (*12–14*), and how well vaccines will protect against more contagious or immune-response-escaping variants. Additional challenges are the potential influx of SARS-CoV-2 infections (brought in by travelers or commuting workers from abroad), imperfect quarantine, limited compliance, and TTI and case-reporting delays. Lastly, any country's capacity to perform TTI is limited, so that spreading dynamics change depending on the level of case numbers. Understanding these dynamics is crucial for informed policy decisions.

Results

Analytical framework: overview

We analytically show the existence of a stable regime at low case numbers, where control of SARS-CoV-2 is much easier to achieve and sustain. In addition, we investigate mitigation strategies and long-term control for COVID-19, where we build on our past work to understand the effectiveness of non-pharmaceutical interventions, particularly test-trace-and-isolate (TTI) strategies (*15–17*). Importantly, the strategy we propose does not rely on the availability of a cure or vaccine, and it is applicable not only to further waves of COVID-19 but also to novel diseases with pandemic potential. Nonetheless, we also show how vaccination campaigns will further facilitate the success of the proposed strategy, assuming a vaccination rate as planned for countries in the European Union (*13*).

For quantitative assessments, we adapt an SEIR-type compartmental model (*18*) to explicitly include a realistic TTI system that considers the challenges above. In our framework, individuals can be tested (and subsequently quarantined if tested positive) by three different mechanisms. First, symptomatic infections with COVID-19-specific symptoms would self-report or be diligently identified by surveillance and get a preferential test. Second, random asymptomatic screening would be homogeneously deployed in the general population disregarding symptoms so that every individual could be tested alike. Third, all the close contacts of those individuals recently tested positive by the two mechanisms mentioned above would also be tested. Understandably, limited resources pose a complex challenge for resource allocation, where efficient TTI would only be possible at low case numbers. We also built an interactive platform where enthusiastic readers can simulate scenarios different from those presented herein `http://covid19-metastability.ds.mpg.de/`.

A central parameter for our analysis is the *level of contagious contacts* k_t (relative to pre-COVID-19). More precisely, k_t refers to the fraction of infection-risk-bearing encounters compared to pre-COVID-19 contact levels. We can then interpret k_t in terms of the hidden reproduction number R_t^H, which accounts for the number of offspring infections generated by individuals unaware of being infectious, i.e., hidden infections in a naive and fully susceptible population (*15*). Thus, in terms of R_t^H, k_t can be understood as the ratio between the offspring infections a hidden individual would generate in the presence and in the absence of NPIs, in other words, $R_t^H = k_t R_0$. Apart from direct contact reduction, contributions that allow increasing k_t without compromising the stability of the system also come from improved hygiene, mandatory face-mask policies, frequent ventilation of closed spaces, and avoiding indoor gatherings, among other precautionary measures. As the latter measures are relatively fixed, direct contact reduction remains the central free variable, which is also the one tuned during lockdowns. All other parameters (and their references) are listed in Table 1.

Equilibrium at low case numbers

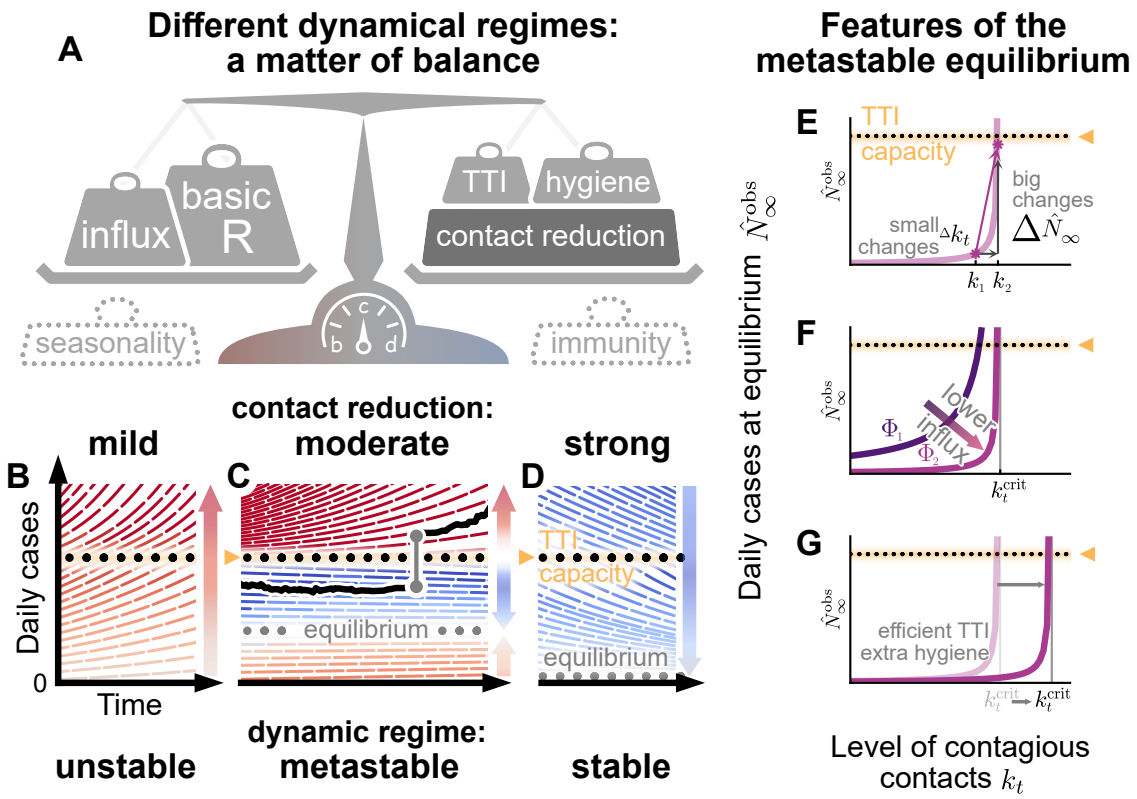


Figure 1: **Spreading dynamics depend on the balance between destabilizing and stabilizing contributions and on the level of case numbers. A:** Among the factors that destabilize the spread, we find the basic reproduction number R_0, and the external influx of infections (and possibly seasonality). On the other hand, increased hygiene, test-trace-isolate (TTI) strategies, contact reduction, and immunity contribute to stability. We specifically investigated how reductions in the contact level k_t and limited TTI capacity determine the stabilization of case numbers. **B:** At mild contact reduction ($k_t = 80\,\%$ compared to pre-COVID-19 times), TTI is not sufficient; case numbers would grow even when TTI capacity is available. **C:** At moderate contact reduction ($k_t = 60\,\%$), a metastable equilibrium emerges (gray dots) to which case numbers converge, if they were below the TTI capacity. However, destabilizing events (e.g., a sudden influx of infections) can push a previously stable system above the TTI capacity and lead to an uncontrolled spread (black line as an example). **D:** Assuming strong contact reduction ($k_t = 40\,\%$), case numbers decrease even if the TTI capacity is exceeded. **E:** Near to the critical level of contacts $k_t^{\rm crit}$, small changes in the contact level will lead to a considerable increase in observed cases in equilibrium $\hat{N}_\infty^{\rm obs}$. **F:** Reducing the influx of infections Φ_t (by closing borders or deploying extensive testing at arrival) reduces the number of infections. **G:** Increasing the efficiency of manual contact tracing and additional measures as increased hygiene and compulsory use of face masks will increase the maximum allowed level of contacts $k_t^{\rm crit}$.

We find a regime where the spread reaches an equilibrium at low daily case numbers between the scenarios of eradication of the disease or uncontrolled spreading. The main control parameter that determines whether the system can reach an equilibrium is the level of contagious contacts k_t.

If the reduction in the contact level k_t is mild, case numbers grow exponentially, as measures could not counterbalance the basic reproduction number ($R_0 \approx 3.3$ for SARS-CoV-2 (*19, 20*)) (Fig. 1B). In contrast, if the reduction in k_t is strong and (together with hygiene and TTI) outweighs the drive by the basic reproduction number, case numbers decrease to a low equilibrium value (Fig. 1D).

Importantly, if the reduction in k_t is moderate (and just-about balances the drive by the basic reproduction number), we find a metastable regime: The spread is stabilized if and only if the overall case numbers are sufficiently low to enable fast and efficient TTI (Fig. 1C). However, this control is lost if the limited TTI

capacity is overwhelmed. Beyond that tipping point, the number of cases starts to grow exponentially as increasingly more infectious individuals remain undetected (*15*).

The capacity of TTI determines the minimal required contact reduction for controlling case numbers around an equilibrium. If case numbers are sufficiently below the TTI capacity limit, the maximum allowed level of contacts to enable the (meta-)stable regime in our default scenario (cf. Table 1, with $R_0 = 3.3$) is $k_t^{\text{crit}} = 61\,\%$ (95% confidence interval (CI):[47, 76]). When the level of contacts k_t is below the threshold, case numbers asymptotically approach to an equilibrium that shows the following features: i) when the level of contacts k_t is close to its critical value k_t^{crit}, small changes in k_t generate large modifications of the equilibrium level (Fig. 1E). ii) Larger influxes lead to larger equilibrium values, however not modifying the maximum allowed level of contacts (Fig. 1F). iii) Behavioral changes and policies leading to a reduction in the transmission probability will also lower the equilibrium value because the maximum allowed level of contacts k_t^{crit} would be larger (Fig. 1G).

However, if case numbers exceed the TTI capacity limit, a considerably stronger reduction in contacts is required to reach the stable regime, so that $k_t^{\text{crit}} = 42\,\%$ (95% CI: [38, 47]) (Fig. 2 B, Fig. S2B and Table S1).

Equilibrium depends on influx and contact reduction

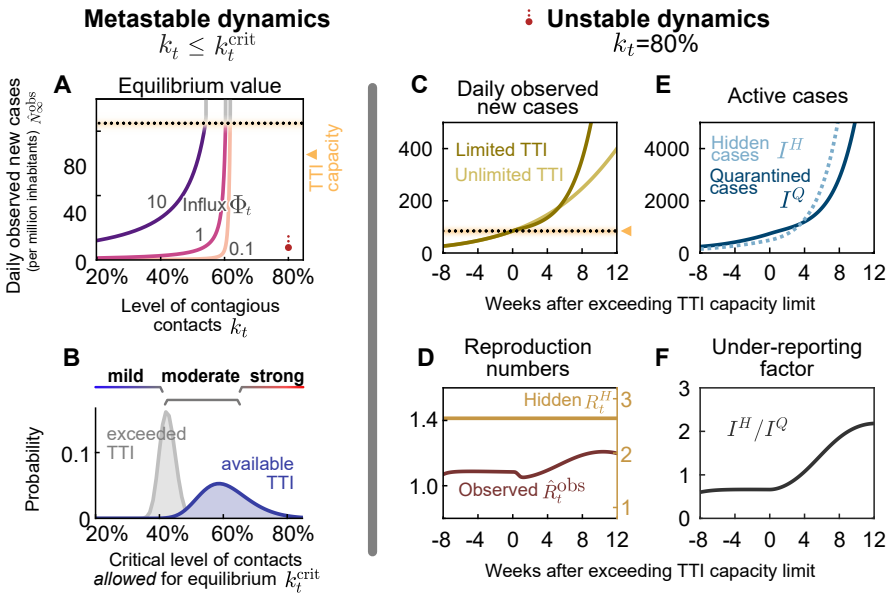


Figure 2: **(A, B) In the stable and metastable regimes, daily new cases approach an equilibrium value $\hat{N}_\infty^{\text{obs}}$ that depends on contact level k_t and the external influx of new cases Φ_t. A:** The equilibrium value $\hat{N}_\infty^{\text{obs}}$ increases with increased contact level k_t or higher influx Φ_t. No equilibrium is reached if either k_t or Φ_t are above the respective (critical) threshold values. **B:** The critical value k_t^{crit} represents the maximal contact level that is allowed to reach equilibrium and stabilize case numbers. If case numbers are below the TTI capacity limit, higher values of k_t^{crit} are permitted for stabilization (blue) than if cases exceed TTI (gray). Confidence intervals originate from error propagation of the uncertainty of the underlying model parameters. **(C–F) In the unstable dynamic regime ($k_t = 80\%$), a tipping point is visible when exceeding TTI capacity.** We observe a self-accelerating increase of case numbers after crossing the TTI limit (**C**) and a subsequent increase of the reproduction number (**D**). Furthermore, the absolute number (**E**) and the fraction (**F**) of cases that remain unnoticed increase over time.

If an equilibrium is reached, the precise value of daily new cases $\hat{N}_\infty^{\text{obs}}$ at which the system stabilizes depends on both the contact level (k_t) and the external influx of new cases (Φ_t) (Fig. 2A). In general, for realistic low values of influx Φ_t, the equilibrium level $\hat{N}_\infty^{\text{obs}}$ is low. However, $\hat{N}_\infty^{\text{obs}}$ increases steeply (diverges) when the contact level k_t approaches the tipping point to unstable dynamics (Fig. 2A,B). Such

a divergence near a critical point k_t^{crit} is a general feature of continuous transitions between stable and unstable dynamics (*21, 22*). As a rule of thumb, in an analytical mean-field approximation (*22*), $\hat{N}_\infty^{\text{obs}}$ would be proportional to Φ_t and diverge when k_t approaches its critical value k_t^{crit} from below:

$$\hat{N}_\infty^{\text{obs}} \propto \frac{\Phi_t}{k_t^{\text{crit}} - k_t}. \tag{1}$$

Robust control of the pandemic requires maintaining a sufficient safety margin from the tipping point (and the subsequent transition to instability) for two reasons. First, small fluctuations in k_t, Φ_t (or other model variables) could easily destabilize the system. Second, near the critical value k_t^{crit}, reductions in k_t are especially effective: already small further reductions below k_t^{crit} lead to significantly lower stable case numbers (Fig. 2A). Already with moderate reductions in k_t ($50\,\% < k_t < 60\,\%$), the spread can be stabilized to a regime of case numbers clearly below 10 per million (Fig. S1B, lower right region).

Limited TTI and self-acceleration

If mitigation measures are insufficient, case numbers rise and eventually surpass the TTI capacity limit. Beyond it, health authorities cannot efficiently trace contacts and uncover infection chains. Thus the control of the spread becomes more difficult. We start our scenario with a slight increase in the case numbers over a few months, as seen in many European countries throughout summer 2020 (Fig. S4 and Fig. S5). A tipping point is then visible in the following observables (Fig. 2 C–F):

First, when case numbers surpass the TTI capacity, the increase in daily new observed cases $\hat{N}^{\text{obs}}$ becomes steeper, growing even faster than the previous exponential growth (Fig. 2 C, full versus faint line). The spread self-accelerates because increasingly more contacts are missed, which, in turn, infect more people. Importantly, in this scenario, the accelerated spread arises solely because of exceeding the TTI limit — without any underlying behavior change among the population.

Second, after case numbers surpass the TTI limit, the observed reproduction number $\hat{R}_t^{\text{obs}}$, which had been only slightly above the critical value of unity, increases significantly by about 20 % (Fig. 2 D). This reflects a gradual loss of control over the spread and explains the faster-than-exponential growth of case numbers. The initial dip in $\hat{R}_t^{\text{obs}}$ is a side-effect of the limited testing: As increasingly many cases are missed, the observed reproduction number reduces transiently.

Third, compared to the infectious individuals who are quarantined I^Q, the number of infectious individuals who are hidden I^H (i.e. those who are not isolated or in quarantine) increases disproportionately (Fig. 2 E) which is measured by the "under-reporting factor" (I^H/I^Q) (Fig. 2 F). The hidden infectious individuals are the silent drivers of the spread as they, unaware of being infectious, inadvertently transmit the virus. This implies a considerable risk, especially for vulnerable people. The TTI system can compensate for the hidden spread at low case numbers because it uncovers hidden cases through contact tracing. However, at high case numbers, the TTI becomes inefficient: If the TTI measures are "slower than the viral spread", many contacts cannot be quarantined before they become spreaders.

Re-establishing control with lockdowns

Once the number of new infections has overwhelmed the TTI system, re-establishing control can be challenging. A recent suggestion is the application of a circuit breaker (*23–25*), a short lockdown to significantly lower the number of daily new infections. Already during the so-called "first wave" of COVID-19 in Europe (i.e., the time-frame between March and June, 2020), lockdowns have proven capable of lowering case numbers by a factor of 2 or more every week (*17, 26*) (corresponding to an observed reproduction number of $\hat{R}_t^{\text{obs}} \approx 0.7$). With the knowledge we now have acquired about the spreading of SARS-CoV-2, more targeted restrictions may yield a similarly strong effect.

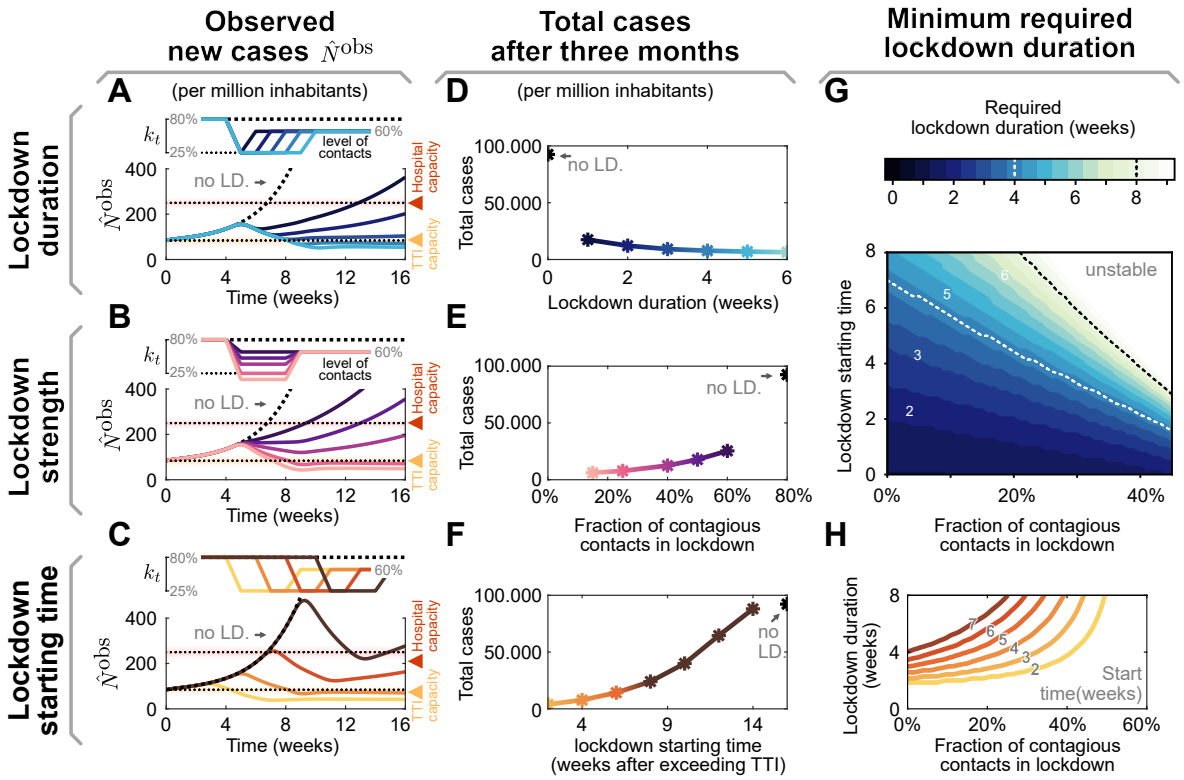


Figure 3: **The effectiveness of a lockdown depends on three primary parameters: its duration, stringency (strength), and starting time.** (**A**–**C**) Observed daily new cases for a lockdown (abbreviated as LD), enacted after exceeding the TTI capacity. Reference parameters are a lockdown duration of 4 weeks, reducing contact level to $k_t = 25\,\%$ and a start time at 4 weeks after exceeding TTI capacity. We vary lockdown duration from 1 to 5 weeks (**A**), lockdown strength (**B**) and lockdown starting time (**C**) to investigate whether stable case numbers can be reached. **D**–**F:** Total cases after three months, if the lockdown is parameterized as described in panels A–C, respectively. **G, H:** The minimal required duration of lockdown to reach equilibrium depends both on strength and start time. **G:** Heavy contact reduction (leading to lower values of k_t) and timely lockdown enacting can create effective short lockdowns (≤ 2 weeks, lower left, dark region). Whereas with mild contact reduction and very late start times, lockdowns become ineffective even when they last indefinitely (upper right, bright area). **H:** Horizontal slices through the color-map (**G**). Here, colors match panels (C, F) and correspond to the lockdown start time.

Inspired by the lockdowns installed in many countries (*27*), we assume a default lockdown of four weeks, starting four weeks after case numbers exceed the TTI capacity limit, and a strong reduction of contagious contacts during a lockdown, leading to $k_t = 25\,\%$ (which corresponds to an $\hat{R}_t^{\text{obs}} \approx 0.85$, see Table S2). We further assume that during lockdown, the external influx of infections Φ_t is reduced by a factor of ten and that after the lockdown, a moderate contact reduction (allowing $k_{\text{nLD}} = 60\,\%$) is maintained. By varying the parameters of this default lockdown, we show in the following that the lockdown strength, duration, and starting time determine whether the lockdown succeeds or fails to reach equilibrium.

In our scenario, a lockdown duration of four weeks is sufficient to reach the stable regime (Fig. 3 A). However, if lifted too early (before completing four weeks), case numbers will rise again shortly afterward. The shorter an insufficient lockdown, the faster case numbers will rise again. Also, it is advantageous to remain in lockdown for a short time even after case numbers have fallen below the TTI limit — to establish a safety margin, as shown above. Overall, the major challenge is not to ease the lockdown too early; otherwise, the earlier success is soon squandered.

During a lockdown, it is necessary to severely reduce contagious contacts to decrease case numbers below the TTI capacity limit (Fig. 3 B). In our scenario, the contact level has to be reduced to at least $k_{\text{LD}} = 25\,\%$

to bring the system back to equilibrium. Otherwise, a lockdown that is slightly weaker would fail to reverse the increasing trend in cases. Furthermore, increasing the lockdown strength decreases both the required lockdown duration (Fig. 3 G, H) and the total number of cases accumulated over three months (Fig. 3 E). This shows that stricter lockdowns imply shorter-lasting social and economic restrictions.

The earlier a lockdown begins after exceeding the TTI capacity limit, the faster control can be re-established, and constraints can be loosened again (Fig. 3 C). If started right after crossing the threshold, in principle, only a few days of lockdown are necessary to bring back case numbers below the TTI capacity limit. On the other hand, if the lockdown is started weeks later, its duration needs to increase (Fig. 3 C,D) and the total number of cases will be significantly larger (Fig. 3 F). The parameter regime between these two options is relatively narrow; it is not likely that equilibrium can eventually be reached as a lockdown exceeds many weeks (cf. Fig. 3 H). For practical policies, this means that if a lockdown does not start to show apparent effects after 2 or 3 weeks, then the strategy should be revised (this assessment time is necessary due to the delay of 1-2 weeks between contagion and case report).

Maintaining control without lockdowns

We show that repeated lockdowns are not required to maintain control over the COVID-19 spread if moderate contact reduction is maintained once case numbers are below the TTI capacity limit (an initial lockdown might still be necessary to establish control). A natural goal would be to keep case numbers below the hospital capacity. However, our model suggests that lowering them below TTI capacity requires fewer contact restrictions (in the long-term), involves a shorter total lockdown duration, and costs fewer lives. In the following, we compare the long-term perspective of these two goals and their dependence on the necessary contact reduction.

In our scenario (Fig. 4), we start from the unstable regime, where the initial contact level ($k_t = 80\,\%$ of the pre-COVID level) is not sufficient to control the spread. We start a two-week lockdown when crossing either the TTI or the hospital capacity. During the lockdown, contacts are reduced to $k_t = 25\,\%$. After the first and all subsequent lockdowns, contacts k_t are reduced to $80\,\%$, $60\,\%$ or $40\,\%$ relative to pre-COVID-19 levels, thus representing a mild, moderate, or strong reduction. When assuming mild contact reduction after lockdowns, case numbers rise after lifting the lockdown, independent of the chosen threshold (TTI or hospital capacity, Fig. 4 A). Thus, repeated lockdowns are necessary.

However, maintaining a moderate contact reduction while not in lockdown (i.e., a contact level of $k_t = 60\,\%$) is sufficient to stay within the metastable regime — if lockdowns are enacted such that case numbers remain below the TTI capacity (Fig. 4 B, yellow line). This is a promising perspective for a long-term control strategy that avoids recurrent lockdowns. Otherwise, if case numbers are above TTI capacity limit but below hospital capacity, control of the pandemic requires repeated lockdowns (Fig. 4 B, red line). Alternatively, lasting strong contact reductions even after the lockdown can be sufficient to drive down case numbers (Fig. 4 C).

The advantage of the strategy to stay below the TTI capacity limit becomes very clear when considering the total cost of the required lockdowns: Independent of the degree of contact reduction, (i) the total number of cases (and consequently deaths and "long-COVID" risk) is lower (Fig. 4 D), (ii) the total duration spent in lockdown is shorter (Fig. 4 E), and (iii) the frequency at which lockdowns have to reoccur – should the after-lockdown contact reduction not be enough to grant metastability – is lower (Fig. 4 A). As case numbers and lockdown duration indicate economic costs, a strategy that respects the TTI limit offers a low economic toll, enables mid-term planning and provides trust to people and society.

The scalability of random testing (screening) (*28–30*) and immunization programs play a critical role in long-term strategies; both will increase the maximum level of contacts allowed (k_t^{crit}) to maintain control (Fig. 4 F, G). Early effects of immunity can be seen in our scenario of the system stabilized at hospital capacity: The need for lockdowns becomes less frequent over time (Fig. 4 A, B red lines). However, acquiring natural

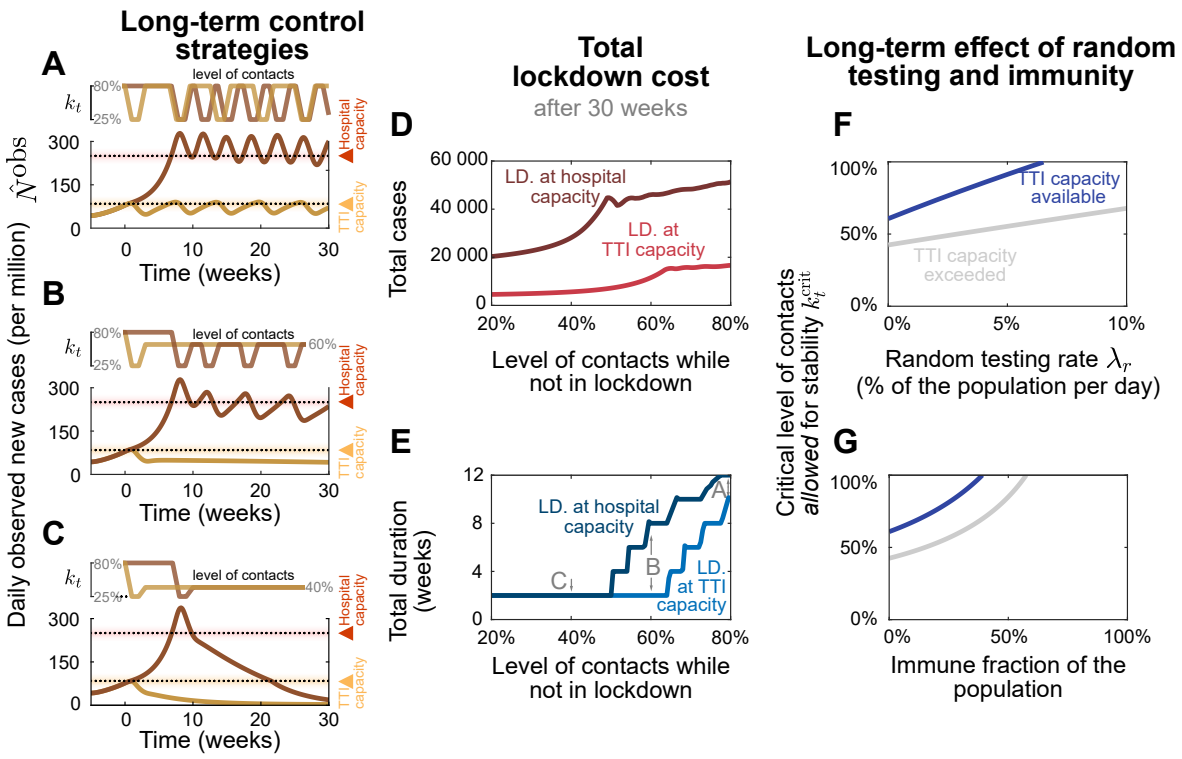


Figure 4: **From a long-term perspective, recurrent lockdowns are not required if the subsequent contact levels k_t are sufficiently low to reach equilibrium.** A two-week lockdown of default strength (reducing contact level to $k_t = 25\,\%$ relative to pre-COVID levels) is either enacted when the test-trace-and-isolate (TTI, yellow lines) or the hospital (red lines) capacity limits are crossed. **A:** Given only a mild contact reduction while not in lockdown (to $k_t = 80\,\%$), the imposed lockdowns need to reoccur with a high frequency. Moreover, the frequency at which they reoccur is higher at high case numbers (aiming at hospital capacity) than when aiming at the TTI capacity limit because of the self-accelerating effect described previously. **B:** If the contact reduction remains moderate (to $k_t = 60\,\%$), lockdowns only need to reoccur if they fail to bring case numbers below TTI capacity. **C:** Only lasting strong contact reduction (to $k_t = 40\,\%$) is sufficient to continue to bring down case numbers after the lockdown. **D, E:** The total cost (in terms of cumulative cases and total lockdown duration) depends on the level of contact reduction and the lockdown policy: The cost is low if lockdowns are initiated at the TTI capacity limit and if the contact level k_t after the lockdown is low. **F, G:** Large-scale random testing (**F**) or immunization of the population (either post-infection or through vaccination **G**) increase the maximal contact level allowed for stability k_t^{crit}.

immunity comes at the cost of a prolonged high level of case numbers, subsequent "long COVID" cases (*4–6*) and deceased people. Whereas the duration of the immunity is yet unknown (*31*), this phenomenon still shows that immunity effects play an increasing role as model predictions extend further into the future.

Vaccination greatly facilitates containment in the effective TTI regime

In the following, we show how growing immunity granted by vaccination programs further facilitates both reaching low case numbers below TTI capacity and the stable control thereof. For quantitative assessments, we studied the effect of COVID-19 vaccination programs and how they affect the two control strategies discussed in the previous section.

Investigating explicit vaccination scenarios, we assume that 80% of people getting offered vaccination accept this offer. Since, as of now, none of the available vaccines has been approved for people younger than 16 years, this corresponds to a vaccine uptake of roughly 70% of the overall population (for European demographics). We model the delivery and administration of all doses to be completed within 32 weeks (*13*),

which is comparable to the increasing vaccination rate in the European Union. We investigate two scenarios: that the average vaccine efficacy against transmission is 80%, in line with reported values for current vaccines against the wild type and the widely dominant B.1.1.7 variant (*32–34*), or 40%, a possible scenario for partially immune-escaping variants.

Comparing the two different control strategies introduced in the preceding section, i.e., either aiming to keep case numbers below the TTI or below the hospital capacity limit, the progressing vaccination Fig. 5 A, B) will eventually lead to declining case numbers Fig. 5 C, D). However, this point will be reached several weeks earlier if case numbers are kept below TTI capacity. Until then, repeated lockdowns will remain necessary if contact reductions outside of lockdown are insufficient (see discussion in the preceding section). A decreasing number of infections will even be reached months earlier if TTI capacity is available than if it is overwhelmed.

Low case numbers are greatly beneficial even in light of progressing vaccination programs. In the high-case-number regime, only in the most optimistic scenario the spread of SARS-CoV-2 can be controlled without contact reductions after the vaccination program is finished (Fig. 5 E, the full gray line reaches 100%). In all other scenarios, because of the dominance of more contagious (dotted lines) or immune escape variants lowering vaccine efficacy (lower row), an efficiently-functioning TTI program is necessary to allow for a high level of contacts (Fig. 5 E, F, blue lines). Therefore reaching low case numbers is complementary to the vaccination efforts and necessary to maximize the population's freedom.

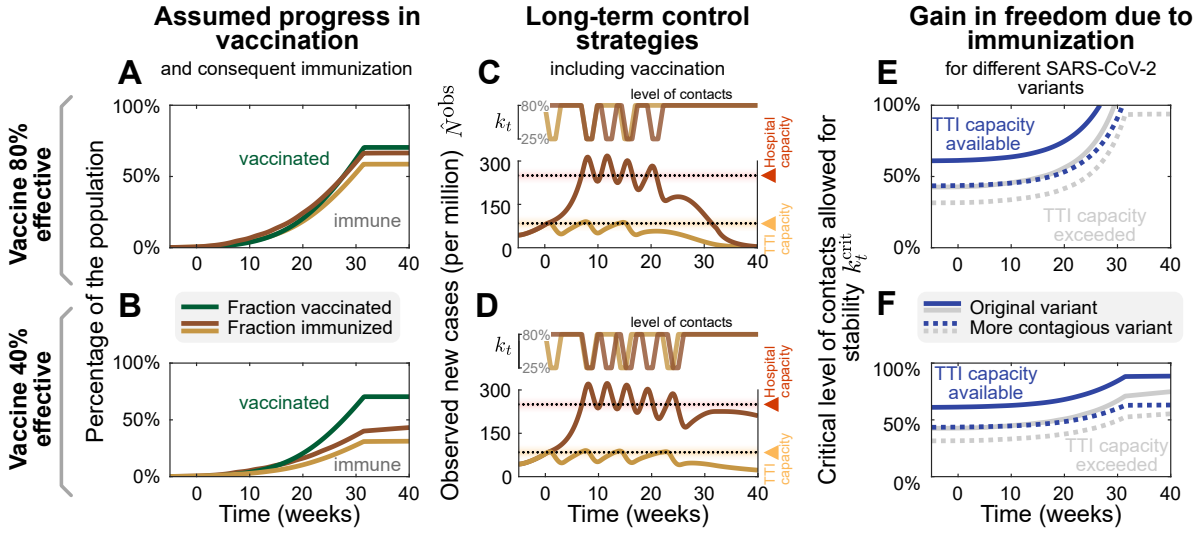


Figure 5: **Growing immunity due to mass vaccination will lead to declining case numbers and enable increasing contact levels** For 80% (upper row) and 40% (lower row) vaccine efficacy against transmission, we explore the effect of mass vaccination. **A, B:** Vaccine rollout is assumed to accelerate over 32 weeks, ending when 70% of the population have been vaccinated. **C, D:** As in Fig. 4 A, a two-week lockdown of default strength (allowing only a $k_t = 25\,\%$ of pre-COVID contacts) is either enacted when the test-trace-and-isolate (TTI, yellow lines) or the hospital (red lines) capacity limits are reached. While not in lockdown, people maintain a level of contacts of $k_t = 80\,\%$. Without a major change in these policies, repeated lockdowns will eventually cease to be necessary due to the growing immunity, and case numbers will decline in all scenarios. However, this decline would occur weeks earlier if case numbers are below TTI capacity and vaccine efficacy is high. **E, F:** Growing immunity among the population will considerably increase the maximal –critical– level of contacts allowed for stabilization of case numbers for both strategies, but more substantially if not exceeding TTI capacity limit. More contagious variants like B.1.1.7 (with an assumed base reproduction number of 4.3) will require a lower level of contacts (dotted lines).

Discussion

We demonstrated that between the two extremes of eradication and uncontrolled spread, a metastable regime of SARS-CoV-2 spreading exists. In such a regime, every person only has to reduce their contacts moderately. Simultaneously, case numbers can still be maintained robustly at low levels because the test-trace-and-isolate (TTI) system can operate efficiently. If this regime is within reach, keeping case numbers below TTI capacity is a suitable strategy for the long-term control of COVID-19 (or other infectious diseases) that features low fatalities and a small societal burden. In addition, it maximizes the effect of large-scale pharmaceutical interventions (as vaccination programs).

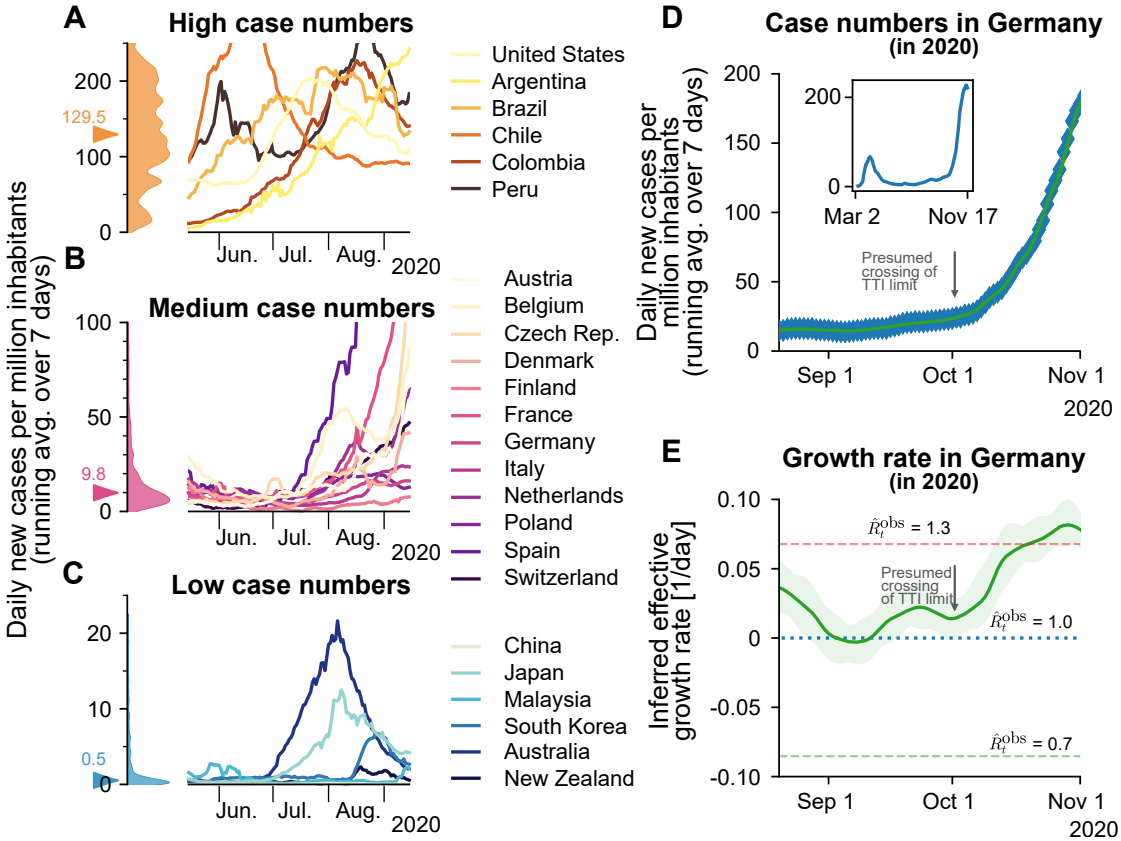


Figure 6: **Strategies of countries to fight SARS-CoV-2 differ widely and are reflected in case numbers. A:** Strategies that involve little non-pharmaceutical interventions rely on the population to hinder the spread in a self-regulated manner and are often accompanied by high case numbers. **B:** Strategies that aim to keep case numbers low through extensive TTI, combined with temporary lockdowns, can lead to medium case numbers. As TTI measures' effectiveness depends on daily infections, case numbers can seemingly explode when the (hard-to-estimate) TTI capacity limit is exceeded. Many European countries managed to stabilize case numbers over summer 2020. However, stability was short-lived, and they quickly saw case numbers rising faster than exponentially. **C:** When the external influx is low or employed strategies to reduce contacts are very effective, the stable regime can be reached. In this case, average case numbers are very low, and local outbreaks can be controlled well through local interventions. Raw data and preliminary visualizations were obtained from (*35*). **D, E:** A substantial increase in Germany's effective growth rate occurred during October, suggesting that regional TTI capacity limits were exceeded. **D:** Observed case numbers were below 20 daily new cases per million until, presumably, a transition into the unstable regime took place over ~ 4 weeks in October. The inset compares the extents of the so-called first and second waves. **E:** Before October, the reproduction number was slightly above one (corresponding to a daily growth rate slightly above zero). More details are in Supplementary Figure S4.

Among countries worldwide, significant variability in governmental policies and the chosen strategy to face COVID-19 significantly impact the levels of observed COVID-19 case numbers (see also Fig. 6 and

Supplementary Notes 1.1). Sustained high levels of more than 100 daily new cases per million have been observed in several (but not exclusively) American countries (Fig. 6A). This shows that high levels of daily new infections can be maintained. However, the stringency of interventions is similar to other countries (*36*), without signs of reaching population immunity. On the other hand, very low case numbers and even local eradication have been achieved by several South- and East-Asian countries, Australia and New Zealand. At the time of writing, these countries profit from the absorbing state at zero SARS-CoV-2 infections, but maintaining this state requires substantial international travel restrictions (Fig. 6C). An intermediate level of case numbers could predominantly be observed over summer 2020 in Europe. Case numbers for many countries were typically around ten daily new cases per million (Fig. 6B), even though contacts were only mildly restricted. These stable numbers demonstrate that also, in practice, a regime below TTI capacity limits is maintainable. Nonetheless, in September, the spread significantly accelerated in several European countries, when case numbers began to exceed 20 to 50 daily new cases per million (Fig. S5). Beyond these levels of case numbers, the TTI systems began to be overwhelmed, making control difficult–in line with our model's results.

In order to focus our model on the general spreading dynamics, we made simplifying assumptions: We assumed that spreading happens homogeneously in the population, with neither regional nor age-related differences. In reality, heterogeneous spreading can lead to regionally differing case numbers, which illustrates the need for regional monitoring of the remaining TTI capacity to allow for early and targeted control measures. In our scenarios, we further assumed that the population's behavior and subsequent contact reduction are constant over time (except during lockdown). Real situations are more dynamic, necessitating frequent reevaluations of the current restrictions and mitigation measures. We also assumed constant TTI effectiveness if below the capacity limit. However, if case numbers are very low, all the available test- and trace efforts could be concentrated on the remaining infection chains. This would further facilitate control at low case numbers. Overall, our analytical results describe the general behavior across countries well and identify the relevant factors for controlling the pandemic.

Quantitatively, our assumptions regarding the efficiency of TTI are in agreement with those of other modeling studies. Agent-based models with detailed contact structures (*37, 38*) and mean-field models (*39–42*) both agree that TTI measures are an essential contribution for the control of the pandemic but typically do not suffice alone. Their success strongly depends on their implementation: Fast testing, rigorous isolation, and a large proportion of traced contacts are essential. Given our informed assumptions about these parameters, our model shows that TTI can only compensate a basic reproduction number R_0 of 3.3 — if contagious contacts are also reduced to at most 61 % (95% CI: [47,76]) compared to pre-COVID. This is in agreement with the results of other studies (*37, 38, 40, 43, 44*).

The capacity limit of TTI plays a central role in the control of the spread but depends strongly on the local environment. The precise limit of TTI depends on several factors, including the number of available tests, personnel at the tracing units, potentially a tracing app (*40*), and the number of relevant contacts a person has on average. Already the latter can easily differ by a factor of 10, depending on contact restrictions and cultural factors (*45*). We assumed the capacity is reached at about 85 daily new cases per million for our scenarios, which is comparably high. Independent of the exact value, when this limit is approached, the risk of tipping over to uncontrolled spread strongly increases, and countermeasures should be taken without delay.

Given the large deviations of the capacity limit of TTI across regions, policymakers should monitor local health authorities' and tracing agencies' capacity instead of relying on fixed limits. Health authorities can also assess whether a local outbreak is controlled or whether infection chains cannot be traced anymore, allowing an early and adaptive warning system. However, one can safely state that daily case numbers larger than 85 per million (corresponding to our modeled limit) are above the capacity limit of TTI programs in Europe, therefore requiring in any case further restrictions to reach controllable levels.

Even given the ongoing vaccination campaigns, a low level of case numbers below TTI capacity limits remains essential. We find that during as well as after the campaigns, TTI still greatly facilitates the containment of COVID-19. The vaccinations' exact effect depends on several hard-to-model factors. It can change with newly-emerging variants of the virus, which can be more contagious, more severe, or escape the immune response. In our analysis of the effect of vaccination, we also neglected age- and high-risk-group distributions and contact networks in the population, the exact design of national vaccination plans, or the differential efficacy of vaccines against infection and severe disease. Some of these factors were taken into account in other publications (*12–14*). However, the long-term success of vaccinations alone remains hard to predict. Thus, it is sensible to accompany mass vaccinations to achieve low case numbers in the vaccine rollout and the time beyond.

Our results show that a stable equilibrium at low case numbers can be maintained with a moderate contact reduction of about 40 % less contagious contacts compared to pre-COVID-19. In terms of our parameters, this translates to a maximum –critical– level of contacts k_t^{crit} of 61 % (95% CI: [47,76]). This level of contacts can be achieved with preventive mitigation measures, as shown by studies analyzing the effectiveness of non-pharmaceutical interventions during the first wave (*17, 26, 27, 46, 47*). Restrictions on the maximum size of gatherings already lead to an effective reduction in the range of 10–40 % (*17, 26, 27, 47*). Improved hygiene, frequent ventilation of rooms, and the compulsory use of masks can further reduce the number of infectious contacts (by a factor that is more difficult to estimate (*48, 49*)). Overall, until mass vaccination plans have been deployed worldwide and available vaccines have been shown to be successful against emerging variants, the regime of low case numbers is very promising for a mid- and long-term management of the pandemic, as it poses the least burden on economy and society.

On the other hand, stabilizing the spread at higher levels of case numbers (e.g. at the hospital capacity limit) requires more stringent and more frequent non-pharmaceutical interventions, because the TTI system cannot operate efficiently. Examples of more stringent measures are the closure of schools and public businesses, stay-at-home orders, and contact ban policies (*17, 26, 46*).

In conclusion, this paper recommends reaching and maintaining low case numbers that allow efficient TTI measures complementary to pharmaceutical interventions. To this end, it is mandatory to counteract local super-spreading events (or an acute influx of infections) as early as possible and to sustain a sufficient level of mitigation measures. If low case numbers are reached and maintained throughout Europe, it will be possible to lift restrictions moderately in the medium-term, and we will be better prepared for the emergence of future variants of concerns.

Methods

Model overview

We model the spreading dynamics of SARS-CoV-2 as the sum of contributions from two pools of infectious individuals, i.e. quarantined-isolated I^Q and hidden non-isolated I^H individuals, while also modeling the infectivity timeline through the incorporation of compartments for individuals exposed to the virus (E^Q, E^H), following an SEIR-like formalism. The quarantined infectious pool (I^Q) contains cases revealed through testing or by contact tracing and subsequently sent to quarantine/isolation to avoid further contacts as well as possible. In contrast, in the hidden infectious pool (I^H), infections spread silently and only become detectable when individuals develop symptoms and get tested, via random testing in the population or as part of the chain of contacts of recently identified individuals. This second pool (I^H) is called the hidden pool; individuals in this pool are assumed to exhibit the general population's behavior, thus of everyone who is not aware of being infected. Healthy individuals that can be infected belong to the susceptible pool S. At the same time, we assume that, after they recover and for the relatively short time frame here studied, they

remain immunized in the R compartment, for a graphical representation of the model, see Fig. 7. We model the mean-field interactions between compartments by transition rates, determining the timescales involved. These transition rates can implicitly incorporate both the disease's time course and the delays inherent to the TTI process. Individuals exposed to the virus become infectious after the latent period, modeled by the transition rate ρ. We distinguish between symptomatic and asymptomatic carriers – this is central when exploring different testing strategies (as detailed below). We also include the effects of non-compliance to TTI measures, modeled as a higher asymptomatic ratio, and imperfect contact tracing, including an explicit delay between testing and contact tracing of contacts. In the different scenarios analyzed, we include a non-zero influx Φ_t of new cases that acquired the virus from outside. Even though this influx makes a complete eradication of SARS-CoV-2 impossible, different outcomes in the spreading dynamics might arise depending on both contact intensity (contact level k_t) and TTI. We then investigate the system's stability and dynamics, aiming to control the spread with a low total number of cases without necessitating a too large reduction of infectious contacts.

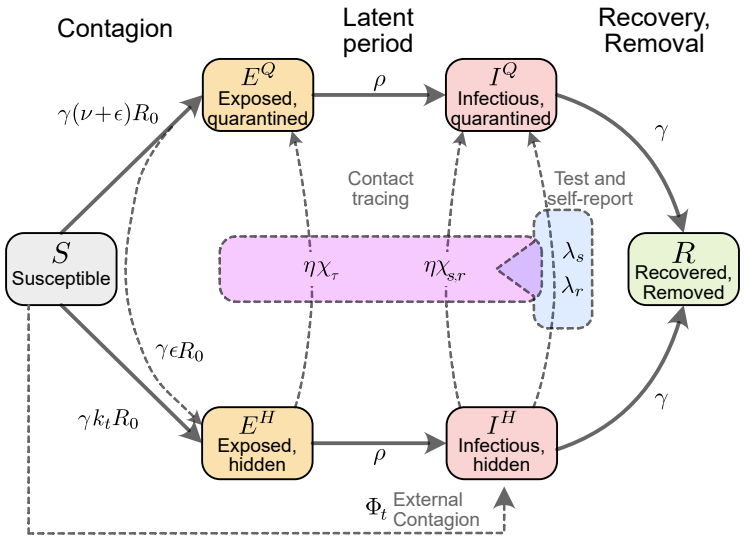


Figure 7: **Flowchart of the complete model.** The solid blocks in the diagram represent different SEIR compartments for both hidden and quarantined individuals. Hidden cases are further divided into symptomatic and asymptomatic carriers (not shown). Solid lines represent the natural progression of the infection (contagion, latent period, and recovery). On the other hand, dashed lines account for imperfect quarantine and limited compliance, external factors, and test-trace-and-isolate policies.

Spreading Dynamics

Concretely, we use a modified SEIR-type model, where infected individuals can be either symptomatic or asymptomatic. They belong to hidden (E^H, I^H) or a quarantined (E^Q, I^Q) pools of infections, thus creating in total one compartment of susceptible (S), two compartments of exposed individuals $(E^H,\ E^Q)$, four compartments of infectious individuals $(I^{H,s},\ I^{H,a},\ I^{Q,s},\ I^{Q,a})$, and one compartment for recovered/removed individuals (R).

New infections are asymptomatic with a ratio ξ, the others are symptomatic. In all compartments, individuals are removed with a rate γ because of recovery or death (see Table 1 for all parameters).

In the hidden pools, the disease spreads according to the population's contact patterns, which can be expressed as a level k_t of the intensity they had before COVID-19 related contact restrictions. Defining R_0 as the base reproduction number without contact restrictions, the reproduction number of the hidden pool I^H is given by $k_t R_0$. This reproduction number reflects the disease spread in the general population without

testing induced isolation of individuals. Additionally, the hidden pool receives a mobility-induced influx Φ_t of new infections. Cases are removed from the hidden pool (i) when detected by TTI and put into the quarantined pool I^Q, or (ii) due to recovery or death.

The quarantined exposed and infectious pools (E^Q, I^Q) contain those infected individuals who have been tested positive as well as their positively tested contacts. Infectious individuals in I^Q are (imperfectly) isolated; we assume their contacts have been reduced to a fraction $(\nu + \epsilon)$ of the ones they had in pre-COVID-19 times, of which only ν are captured by the tracing efforts of the health authorities. The subsequent infections remain quarantined, thus entering the E^Q pool and afterward the I^Q pool. The remaining fraction of produced infections, ϵ, are missed and act as an influx to the hidden pools (E^H). Therefore, the overall reproduction number in the I^Q pool is $(\nu + \epsilon)\, R_0$.

As our model is an expanded SEIR model, it assumes post-infection immunity, which is a realistic assumption given the limited time frame considered in our analysis. Our model can also reflect innate immunity; one has to rescale the population or the reproduction number. The qualitative behavior of the dynamics is not expected to change.

Parameter Choices and Scenarios

For any testing strategy, the fraction of infections that do not develop any symptoms across the whole infection timeline is an important parameter, and this also holds for testing strategies applied to the case of SARS-CoV-2. In our model, this parameter is called ξ^{ap} and includes beside true asymptomatic infections ξ also the effect of individuals that avoid testing (*50*). The exact value of the fraction of asymptomatic infections ξ, however, is still fraught with uncertainty, and it also depends on age (*51*). While early estimates were as high as 50 % (for example ranging from 26 % to 63 % (*52*)), these early estimates suffered from reporting bias, small sample sizes and sometimes included pre-symptomatic cases as well (*53*). Recent studies estimate the asymptomatic transmission to be more minor (*54*), estimates of the fraction of asymptomatic carriers range between 12 % (*53*) and 33 % (*55*).

Another crucial parameter for any TTI strategy is the reproduction number of the hidden infections. This parameter is by definition impossible to measure, but it is typically the main driver of the spreading dynamics. It depends mainly on the contact behavior of the population and ranges from R_0 in the absence of contact restrictions to values below 1 during strict lockdown (*17*). Herein we decided to include instead contact level compared to the pre-COVID-19 baseline k_t to represent the reproduction number of hidden infections $R_t^H = k_t R_0$. For the default parameters of our model, we evaluated different contact levels k_t.

Testing-and-Tracing strategies

We consider a testing-and-tracing strategy: symptom-driven testing and specific testing of traced contacts, with subsequent isolation (quarantine) of those who tested positive. Our model can also include random testing, but this case is only explored in Fig. 4 of this paper.

Symptom-driven testing is defined as applying tests to individuals presenting symptoms of COVID-19. In this context, it is important to note that non-infected individuals can have symptoms similar to those of COVID-19, as many symptoms are rather unspecific. Although symptom-driven testing suffers less from imperfect specificity, it can only uncover symptomatic cases that are willing to be tested (see below). Here, *symptomatic, infectious individuals* are transferred from the hidden to the traced pool at rate λ_s.

We define λ_s as the daily rate at which symptomatic individuals get tested among the subset who are willing to get tested because of surveillance programs or self-report. As default value, we use $\lambda_s = 0.25$, which means that, on average, an individual willing to get tested that develops COVID-19-specific symptoms would get a test within four days from the end of the latency period. Testing and isolation happen immediately in

this model, but their report into the observed new daily cases $\hat{N}^{\text{obs}}$ is delayed, and so is the tracing of their contacts.

Tracing contacts of positively tested infectious individuals presents a very specific test strategy and is expected to be effective in breaking the infection chains if contacts self-isolate sufficiently quickly (*37, 56, 57*). However, as every implementation of a TTI strategy is bound to be imperfect, we assume that only a fraction $\eta < 1$ of all contacts can be traced. These contacts, if tested positive, are then transferred from the hidden to the quarantined infectious pools ($I^H \to I^Q$) with an average delay of $\tau = 2$ days. The parameter η effectively represents the fraction of secondary and tertiary infections that are found through contact tracing. As this fraction decreases when the delay between testing and contact-tracing increases, we assumed a default value of $\eta = 0.66$, i.e. on average, only two-thirds of subsequent offspring infections are prevented. Contact tracing is mainly done manually by the health authorities in Germany. This limits the maximum number $N^{\text{test}}_{\text{max}}$ of new cases observed through testing N^{test}, for which contact tracing is still functional.

Random testing is defined here as applying tests to individuals irrespective of their symptom status or whether they belonged to the contact chain of other infected individuals. In our model, random testing transfers infected individuals from the hidden to the quarantined infectious pools with a fixed rate λ_r, irrespective of they are showing symptoms or not. In reality, random testing is often implemented as situation-based testing for a sub-group of the population, e.g. at a hotspot, for groups at risk, or for people returning from travel. Such situation-based strategies would be more efficient than the random testing assumed in this model, which may be unfeasible at a country level due to testing limitations (*15*).

Lockdown modeling

To assess the effectiveness of lockdowns in the broad spectrum of contact-ban governmental interventions, we model how the reduction of contacts and the duration of such restrictive regimes help lower case numbers.

We model contact reductions as reductions in the reproduction number of the hidden population, which for these matters is presented as percentages of the basic reproduction number R_0, which sets the pre-COVID-19 baseline for the number of close contacts.

For the sake of simplicity, we assume the lockdown scenarios have three stages: i) an uncontrolled regime, where the TTI capacity is overwhelmed because of high case numbers and unsustainable contact levels, reflected by a high value of k_t and a high influx of infections Φ_t. ii) Lockdown is enacted, imposing a strong reduction of contacts, leading to lower values of k_t, and borders closing, leading to a lower influx of Φ_t. iii) Measures are relaxed, allowing higher levels of contacts k_t and restoring international transit. All the changes between the different regimes i→ii→iii are modeled as linear ramps for both parameters, which take $D_{\text{ramp}} = 7$ days to reach their set-point. The duration of the lockdown, namely, the time frame between the start of the restrictive measures and the beginning of their relaxation, is measured in weeks. Its default length –for analysis purposes– is $D_{\text{L}} = 4$ weeks. These values have been chosen following the results of (*27*), where the first effects of an NPI were seen after seven days, and the maximum effect after four weeks.

Model Equations

The contributions of the spreading dynamics and the TTI strategies are summarized in the equations below. They govern the dynamics of case numbers between the different susceptible-exposed-infectious-recovered (SEIR) pools, both hidden (non-isolated) and quarantined. We assume a regime where most of the population is susceptible, and the time frame analyzed is short enough to assume post-infection immunity. Thus, the dynamics are completely determined by the spread (characterized by the reproduction numbers $k_t R_0$ and $(\nu + \epsilon)\, R_0$), the transition from exposed to infectious (at rate ρ), recovery (characterized by the recovery rate γ), external influx Φ_t and the impact of the TTI strategies:

$$\frac{dS}{dt} = -\underbrace{\gamma k_t R_0 \frac{S}{M} I^H}_{\text{hidden contagion}} - \underbrace{\gamma(\nu+\epsilon) R_0 \frac{S}{M} I^Q}_{\text{traced contagion}} - \underbrace{\frac{S}{M}\Phi_t}_{\text{ext. influx}}, \tag{2}$$

$$\frac{dE^Q}{dt} = \underbrace{\gamma \nu R_0 \frac{S}{M} I^Q}_{\text{traced contagion}} + \underbrace{\chi_\tau N^{\text{traced}}}_{\text{contact tracing}} - \underbrace{\rho E^Q}_{\text{end of latency}}, \tag{3}$$

$$\frac{dE^H}{dt} = \underbrace{\gamma \frac{S}{M}\left(k_t R_0 I^H + \epsilon R_0 I^Q\right)}_{\text{hidden contagion}} - \underbrace{\chi_\tau N^{\text{traced}}}_{\text{contact tracing}} - \underbrace{\rho E^H}_{\text{end of latency}}, \tag{4}$$

$$\frac{dI^Q}{dt} = \underbrace{\rho E^Q - \gamma I^Q}_{\text{spreading dynamics}} + \underbrace{N^{\text{test}}}_{\text{testing}} + \underbrace{\left(\chi_{s,r}(1-\xi) + \chi_r \xi\right) N^{\text{traced}}}_{\text{contact tracing}}, \tag{5}$$

$$\frac{dI^H}{dt} = \underbrace{\rho E^H - \gamma I^H}_{\text{spreading dynamics}} - \underbrace{N^{\text{test}}}_{\text{testing}} - \underbrace{\left(\chi_{s,r}(1-\xi) + \chi_r \xi\right) N^{\text{traced}}}_{\text{contact tracing}} + \underbrace{\frac{S}{M}\Phi_t}_{\text{ext. influx}}, \tag{6}$$

$$\frac{dI^{H,s}}{dt} = \underbrace{(1-\xi)\rho E^H - \gamma I^{H,s}}_{\text{spreading dynamics}} - \underbrace{N_s^{\text{test}}}_{\text{testing}} - (1-\xi)\left(\underbrace{\chi_{s,r} N^{\text{traced}}}_{\text{contact tracing}} - \underbrace{\frac{S}{M}\Phi_t}_{\text{ext. influx}}\right), \tag{7}$$

$$I^{H,a} = I^H - I^{H,s}, \tag{8}$$

$$\frac{dR}{dt} = \underbrace{\gamma\left(I^Q + I^H\right)}_{\text{recovered/removed individuals}}. \tag{9}$$

Initial conditions

Let x be the vector collecting the variables of all different pools:

$$x = [S,\, E^Q,\, E^H,\, I^Q,\, I^H,\, I^{H,s},\, R]. \tag{10}$$

We assume a population size of $M = 10^6$ individuals, so that $\sum_{i\neq 6} x_i = M$, and a prevalence of $I_0 = 200$ infections per million, so that $I^Q(0) = I_0$. Assuming that the hidden amount of infections is in the same order of magnitude I_0, we would have $I^H(0) = I_0$, $I^{H,s}(0) = (1-\xi)I_0$. We would expect the exposed individuals to scale with $k_t R_0 I_0$, but we rather assume them to have the same size of the corresponding infectious pool. To calculate the initially susceptible individuals, we use $S(0) = 1 - \sum_{i\neq\{1,6\}} x_i$.

Effect of delays and capacity limit on the effectiveness of TTI strategies

In this section we discuss further details on the derivation of the different parameters and variables involved in equations (2)–(9). First, as we assume contact tracing to be effective after a delay of τ days, some of the individuals who acquired the infection from those recently tested might have also become infectious by the time of tracing. Moreover, a fraction of those who became infectious might also have been tested by the tracing time, should they have developed symptoms.

Furthermore, we give explicit forms for N^{test} and N^{traced} the number of cases identified respectively by testing and contact-tracing. When surpassing TTI capacity, we assume that both testing and contact-tracing change their dynamics simultaneously. This happens when the daily amount of cases identified by testing N^{test} overpasses the TTI threshold $N^{\text{test}}_{\text{max}}$. After being overwhelmed, the overhead testing would change its rate $\lambda_s \to \lambda'_s$, as only patients with a more specific set of symptoms would be tested. Nonetheless, the

contact-tracing efforts can only follow the contacts of those $N_{\max}^{\text{test}}$ observed cases, identifying a fraction η of the offspring infections they produced in their infectious period spent in unawareness of their state. The possibility of random testing is analyzed in Supplementary Note 1.6.

Limited testing capacity leading to lower testing rates

In the first stages of an outbreak, individuals with any symptoms from the broad spectrum of COVID-19-related symptoms would be tested, disregarding how specific those symptoms are. At this stage, we assume that the rate at which symptomatic individuals are tested is λ_s, such that the number of individuals identified through testing (which, for simplicity, is assumed to be solely symptom-driven, i.e., $\lambda_r = 0$) is given by

$$N^{\text{test}} = N_s^{\text{test}} = \lambda_s I^{H,s}. \tag{11}$$

If in addition some random testing, independent on symptomatic status, is performed ($\lambda_r \neq 0$), then $N^{\text{test}} \neq N_s^{\text{test}}$. For this case see S1.6.

When reaching the daily number $N_{\max}^{\text{test}}$ of positive tests, the testing capacity is reached. We then assume that further tests are only carried out for a more specific set of symptoms, leading to a smaller fraction of the tested population. We, therefore, implement the testing capacity as a soft threshold. Assuming that after reaching $N_{\max}^{\text{test}}$, the testing rate for further cases would decrease to λ_s', the testing term N^{test} would be given by

$$N^{\text{test}} = \lambda_s \min\left(I^{H,s}, I_{\max}^{H,s}\right) + \lambda_s' \max\left(0, I^{H,s} - I_{\max}^{H,s}\right), \tag{12}$$

where $I_{\max}^{H,s}$ represent the size of the infectious -symptomatic, hidden- pool i.e, $I_{\max}^{H,s} = \frac{N_{\max}^{\text{test}}}{\lambda_s}$.

Modeling the number of traced individuals

To calculate the number of traced individuals, we assume that a fraction η of the newly tested individuals' contacts, and therefore their offspring infections, will be traced and subsequently quarantined. However, in the presence of TTI, individuals stay on average a shorter amount of time in the infectious pool because they are quarantined before recovering. Therefore, the number of offspring infections has to be corrected by a factor, the average residence time in the infectious pool. For the case $\lambda_r = 0$, the average residence time is $\frac{1}{\gamma+\lambda_s}$, as $I^{H,s}$ is emptied by $-\gamma I^{H,s} - N_s^{\text{test}} = -(\gamma + \lambda_s) I^{H,s}$, i.e., with a rate $\gamma + \lambda_s$. The average residence time in the absence of TTI (natural progression of the disease) is $\frac{1}{\gamma}$. Dividing these two times gives us the wanted correction factor. Thus, the number of traced persons N^{traced} at time t is a fraction η from the offspring infections generated during the residence time, per each individual:

$$N^{\text{traced}}(t) = \eta R_{t-\tau} \frac{\gamma}{\gamma + \lambda_s} N^{\text{test}}(t - \tau), \tag{13}$$

where $R_{t-\tau}$ represents the effective reproduction number:

$$R_{t-\tau} = k_{t-\tau} R_0 \frac{S}{M}. \tag{14}$$

In other words, the number of infectious individuals found by contact tracing at time t are a fraction η of the number of offspring infections generated by individual while they were untested $R_{t-\tau}\frac{\gamma}{\gamma+\lambda_s}$, times the number of individuals tested τ days ago $N^{\text{test}}(t - \tau)$. However, when the TTI capacity is overwhelmed, we assume that the number of traced individuals is limited, that only the contacts of $N_{\max}^{\text{test}}$ individuals (already introduced in the previous section) can be traced:

$$N^{\text{traced}}(t) = \eta R_{t-\tau} \frac{\gamma}{\gamma + \lambda_s} N_{\max}^{\text{test}}. \tag{15}$$

Individuals becoming infectious or being tested by the time of tracing

The traced individuals are removed from either the exposed hidden pool E^H or from the infectious hidden pool I^H after a delay of τ days after testing. As we assume a tracing delay τ of only two days, a significant fraction of the traced individuals would still be in exposed compartments by the time of contact tracing. However, some might already become infectious by that time. To calculate the exact fraction of individuals remaining in the hidden exposed pool by the time of tracing, we proceed as follows. Let $s \in I_\tau = [0, \tau]$ be the time elapsed from the moment of testing. The emptying of the normalized exposed compartment (denoted $\widetilde{E^H}$) due to progression to the infectious stage follows first-order kinetics:

$$\frac{d\widetilde{E^H}}{ds} = -\rho\widetilde{E^H}, \qquad \widetilde{E^H}(0) = 1 \tag{16}$$

The solution of (16) is given by $\widetilde{E^H(s)} = \exp(-\rho s)$. Therefore, we define χ_τ as the fraction of the traced individuals remaining in the E^H compartment at $s = \tau$:

$$\chi_\tau = \exp(-\rho\tau). \tag{17}$$

The remaining individuals are removed from the infectious compartment, which are then simply described by the fraction

$$\chi_r = 1 - \chi_\tau. \tag{18}$$

This, however, only holds for the asymptomatic hidden infectious pool. For the symptomatic hidden pool $I^{H,s}$, we do not want to remove the individuals who have already been tested, as they would be removed twice. For modeling the fraction of non-tested individuals remaining in the normalized symptomatic infectious compartment (denoted $\widetilde{I^{H,s}}$), we couple two first-order kinetics:

$$\frac{d\widetilde{I^{H,s}}}{ds} = -\lambda_s\,\widetilde{I^{H,s}} + \rho\widetilde{E^H}, \qquad \widetilde{I^{H,s}}(0) = 0. \tag{19}$$

The solution of (19) depends on whether $\lambda_s = \rho$ or not. The solution at $s = \tau$ which is the fraction of traced individuals removed from $I^{H,s}$ is given by:

$$\chi_{s,r} = \begin{cases} \rho\tau\exp(-\rho\tau) & \text{if } \lambda_s \approx \rho, \\ \dfrac{\rho}{\lambda_s - \rho}\left(\exp(-\rho\tau) - \exp(-\lambda_s\tau)\right) & \text{else.} \end{cases} \tag{20}$$

For the case $\lambda_r \neq 0$, the reader is referred to the Supplementary Note 1.6.

Including the effects of ongoing vaccination campaigns

To incorporate the effects of COVID-19 vaccination programs in our model, we made some simplifying assumptions. First, we assume that vaccinated individuals have a probability κ of not being infected even if they have a contact with somebody infectious, then not contributing to the spreading dynamics. We define this parameter as the "vaccine efficacy against infection", which has been reported around 50–90% for available vaccines (*32–34*). Thus, they can be assumed to have developed perfect immunity and therefore can be removed from the susceptible (S) and put into the removed compartment (R). The assumption above also implies that these individuals would not take part in the TTI scheme, which would resemble the growing trend of "vaccination passports". Second, we assume a logistically increasing daily vaccination rate $v(t)$ consistent with the projections in the European Union (EU) and assume that 70% of the total population get vaccinated

(see Fig. 5 B, E). This would, e.g., in Germany, amount to roughly 80 % of the adult population (16+ years old) accepting the offer of vaccination, since as of now, none of the available vaccines is approved for children. We find that efficient TTI can significantly enhance the effect of the growing immunity.

To include gradually growing immunity due to an ongoing vaccination campaign we modify (2) with an additional term $-\kappa \cdot v(t)$, as well as (9) with a $+\kappa \cdot v(t)$, with a daily vaccination rate

$$v(t) = \frac{9.3 \times 10^3}{1 + \exp\left(-0.025(t - 150)\right)} \text{ doses per million per day,} \tag{21}$$

centered at $t = 0$, which denotes the start of the vaccinations. This logistic increase in vaccination rates and parameters involved were adapted from (*13*) and roughly mirrors the projected vaccinations in the EU (projections dated to the beginning of February 2021). The factor 9.3×10^3 is determined assuming that after $t_{\text{ref}} = 220$ days, 70% of the population would be vaccinated. Using as reference the age distribution of Germany, the above would amount to roughly a 80 % of the adult population (16+ years old) accepting the offer of vaccination. After that time, for simplicity, we assume the vaccination stops (and therefore $v(t) = 0$, for $t > 220$).

This treatment of the vaccination is simplistic. In reality, most currently available vaccines imply receiving two doses in the span of a few weeks, where the first only gives partial protection. Furthermore, vaccinated individuals need some time to develop a proper immune response after receiving the vaccine (*58*), in which they can still get infected. We also do not incorporate the efficacy of vaccines against a severe course of the disease or death. Since vaccinated but yet infected individuals would have a lower chance of being admitted to the hospital, this would falsify our assumption that hospital capacity can be adequately measured by case numbers alone. As more and more of the daily new infections correspond to individuals already vaccinated, hospitals would only fill up at higher case numbers. To include this effect, the distribution of high-risk groups in the population and the prioritized vaccination programs would have to be taken into account. Including all this is beyond the scope of this work. We addressed in a separate work, building on the results presented herein (*13*). Yet, this simplified implementation is sufficient for our qualitative assessments.

Central epidemiological parameters that can be observed

In the real world, the disease spread can only be observed through testing and contact tracing. While the *true* number of daily infections N is a sum of all new infections in the hidden and traced pools, the *observed* number of daily infections $\hat{N}^{\text{obs}}$ is the number of new infections discovered by testing, tracing, and surveillance of the contacts of those individuals in the quarantined infectious pool I^Q, delayed by a variable reporting time. This includes internal contributions and contributions from testing and tracing:

$$N = \underbrace{\gamma k_t R_0 \frac{S}{M} I^H}_{\text{hidden contagion}} + \underbrace{\gamma\left(\nu + \epsilon\right) R_0 \frac{S}{M} I^Q}_{\text{traced contagion}} + \underbrace{\frac{S}{M}\Phi_t}_{\text{ext. influx}} \tag{22}$$

$$\hat{N}^{\text{obs}} = \Big[\underbrace{\rho E^Q}_{\text{traced contagion}} + \underbrace{N^{\text{test}} + \left(\chi_{s,r}\left(1-\xi\right) + \chi_r \xi\right) N^{\text{traced}}}_{\text{TTI}} \Big] \circledast \mathcal{K}, \tag{23}$$

where $\circledast$ denotes a convolution and $\mathcal{K}$ an empirical probability mass function that models a variable reporting delay, inferred from German data (as the RKI reports the date the test is performed, the delay until the appearance in the database can be inferred): The total delay between testing and reporting a test corresponds to one day more than the expected time the laboratory takes for obtaining results, which is defined as follows: from testing, 50 % of the samples would be reported the next day, 30 % the second day, 10 % the third day, and further delays complete the remaining 10 %, which for simplicity we will truncate at day

four. Considering the extra day needed for reporting, the probability mass function for days 0 to 5 would be given by $\mathcal{K} = [0, 0, 0.5, 0.3, 0.1, 0.1]$. The spreading dynamics are usually characterized by the observed reproduction number $\hat{R}_t^{\text{obs}}$, which is calculated from the observed number of new cases $\hat{N}^{\text{obs}}(t)$. We here use the definition underlying the estimates that Robert-Koch-Institute publishes, the official body responsible for epidemiological control in Germany (*59*): the reproduction number is the relative change of daily new cases N separated by four days (the assumed serial interval of COVID-19):

$$\hat{R}_t^{\text{obs}} = \frac{\hat{N}^{\text{obs}}(t)}{\hat{N}^{\text{obs}}(t-4)} \tag{24}$$

In contrast to the original definition of $\hat{R}_t^{\text{obs}}$ (*59*), we do not need to remove real-world noise effects by smoothing this ratio.

Numerical calculation of solutions and critical values.

The numerical solution of the delay differential equations governing our model were obtained using a versatile solver that tracks discontinuities and integrates with the explicit Runge-Kutta (2,3) pair, `@dde23` implemented in MATLAB (version 2020a), with default settings. This algorithm allows the solution of non-stiff systems of differential equations in the shape $y'(t) = f(t, y(t), y(t-\tau_1), ..., y(t-\tau_k))$, for a set of discrete lags $\{\tau_i\}_{i=1}^k$. Suitability and details on the algorithm are further discussed in (*60*).

To derive the tipping point between controlled and uncontrolled outbreaks (e.g., critical, minimal required contact reduction k_t^{crit} for both stability and metastability), and to plot the stability diagrams, we used the `@fzero` MATLAB function, and the linear approximation of the system of DDE (3)–(7) for the $\frac{S}{M} \approx 1$ limit. This function uses a combination of bisection, secant, and inverse quadratic interpolation methods to find the roots of a function. For instance, following the discussion of Supplementary Notes 1.2, the different critical values for the contact reduction k_t^{crit} were determined by systematically solving the nonlinear eigenvalues problem for stability (*61*), where the solution operation was approximated with a Chebyshev differentiation matrix (*62*).

We also study the effect of dividing the exposed compartment into three sub-compartments, thereby reducing the variability of the latent period distribution (understood as the distribution of waiting times from being infected until becoming infectious). We explored this extended system's linear stability in Supplementary Note 1.8 and confirmed that using a single compartment efficiently characterizes the tipping points.

Table 1: Model parameters.

Parameter	Meaning	Value (default)	Range	Units	Source
M	Population size	1 000 000		people	-
R_0	Basic reproduction number	3.3	2.2–4.4	–	(*19, 20*)
ν	Registered contacts (quarantined)	0.075		–	Assumed
ϵ	Lost contacts (quarantined)	0.05		–	Assumed
γ	Recovery/removal rate	0.10	0.08–0.12	day^{-1}	(*63, 64*)
ξ	Asymptomatic ratio	0.32	0.15–0.43	–	(*52, 53, 55*)
λ_s	Symptom-driven testing rate	0.25	0–1	day^{-1}	Assumed
λ'_s	Symptom-driven testing rate (reduced capacity)	0.1		day^{-1}	Assumed
λ_r	Random testing rate	0.0	0.0–0.1	day^{-1}	Assumed
η	Tracing efficiency	0.66		–	Assumed
τ	Contact tracing delay	2		days	Assumed
$N_{\text{max}}^{\text{test}}$	Maximal tracing capacity	50	10–75	cases day^{-1}	Assumed
Φ_t	External influx	1		cases day^{-1}	Assumed
ρ	Exposed-to-infectious rate	0.25		day^{-1}	(*10, 65*)
D_{L}	lockdown duration	4	0-8	weeks	(*27*)
D_{ramp}	Phase-transition duration (lockdown)	1		weeks	(*27*)
χ_τ	Fraction of contacts traced before becoming infectious	0.61		–	eq (17)
$\chi_{s,r}$	Fraction of contacts traced after becoming infectious, before being tested (symptomatic and random)	0.30		–	eq (20)
χ_r	Fraction of contacts traced after becoming infectious, before being tested (random)	0.39		–	eq (18)

Table 2: Model variables.

Variable	Meaning	Units	Explanation
S	Susceptible pool	people	non-infected people that may acquire the virus.
E^Q	Exposed pool (quarantined)	people	Total quarantined exposed people.
E^H	Exposed pool (hidden)	people	Total non-traced, non-quarantined exposed people.
$I^{H,s}$	Infectious pool (hidden, symptomatic)	people	Non-traced, non-quarantined people who are symptomatic.
I^H	Infectious pool (hidden)	people	Total non-traced, non-quarantined infectious people.
I^Q	Infectious pool (quarantined)	people	Total quarantined infectious people.
N	New infections (Total)	$\text{cases}\,\text{day}^{-1}$	Given by: $N = \gamma k_t R_0 I^H + \gamma\left(\nu + \epsilon\right) R_0 I^Q + \frac{S}{M}\Phi_t$.
k_t	Contact reduction	%	Reduction of infectious contacts, related to pre-COVID-19 times.
$\hat{N}^{\text{obs}}$	Observed new infections (influx to traced pool)	$\text{cases}\,\text{day}^{-1}$	Daily new cases, observed from the quarantined pool; delayed because of imperfect reporting and realistic contact tracing.
$\hat{R}_t^{\text{obs}}$	Observed reproduction number	–	The reproduction number that can be estimated only from the observed cases: $\hat{R}_t^{\text{obs}} = \hat{N}^{\text{obs}}(t)/\hat{N}^{\text{obs}}(t-4)$.
N^{test}	Number of cases found through testing	people	Cases can be found either through symptomatic or random testing $N^{\text{test}} = N_r^{\text{test}} + N_s^{\text{test}}$.
N^{traced}	Number of uncovered infections through tracing	people	This number is limited (depending on the reproduction number) by the maximal tracing capacity $N_{\text{max}}^{\text{test}}$

References

[1] H. E. Randolph, L. B. Barreiro, Herd immunity: Understanding covid-19. *Immunity* **52**, 737–741 (2020).

[2] A. T. Levin, W. P. Hanage, N. Owusu-Boaitey, K. B. Cochran, S. P. Walsh, G. Meyerowitz-Katz, Assessing the age specificity of infection fatality rates for covid-19: systematic review, meta-analysis, and public policy implications. *European journal of epidemiology* pp. 1–16 (2020).

[3] N. A. Alwan, R. A. Burgess, S. Ashworth, R. Beale, N. Bhadelia, D. Bogaert, J. Dowd, I. Eckerle, L. R. Goldman, T. Greenhalgh, D. Gurdasani, A. Hamdy, W. P. Hanage, E. B. Hodcroft, Z. Hyde, P. Kellam, M. Kelly-Irving, F. Krammer, M. Lipsitch, A. McNally, M. McKee, A. Nouri, D. Pimenta, V. Priesemann, H. Rutter, J. Silver, D. Sridhar, C. Swanton, R. P. Walensky, G. Yamey, H. Ziauddeen, Scientific consensus on the COVID-19 pandemic: we need to act now. *The Lancet* **6736**, 19–20 (2020).

[4] E. Fraser, Long term respiratory complications of covid-19. *BMJ* **370** (2020).

[5] T. Greenhalgh, M. Knight, C. A Ćourt, M. Buxton, L. Husain, Management of post-acute COVID-19 in primary care. *BMJ* **370** (2020).

[6] E. J. Topol, COVID-19 can affect the heart. *Science* **370**, 408–409 (2020).

[7] V. Priesemann, M. M. Brinkmann, S. Ciesek, S. Cuschieri, T. Czypionka, G. Giordano, D. Gurdasani, C. Hanson, N. Hens, E. Iftekhar, M. Kelly-Irving, P. Klimek, M. Kretzschmar, A. Peichl, M. Perc, F. Sannino, E. Schernhammer, A. Schmidt, A. Staines, E. Szczurek, Calling for pan-European commitment for rapid and sustained reduction in SARS-CoV-2 infections. *The Lancet* (2020).

[8] A. Scherbina, Determining the optimal duration of the COVID-19 suppression policy: A cost-benefit analysis. *Available at SSRN 3562053* (2020).

[9] J. Xiong, O. Lipsitz, F. Nasri, L. M. Lui, H. Gill, L. Phan, D. Chen-Li, M. Iacobucci, R. Ho, A. Majeed, R. S. McIntyre, Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of Affective Disorders* **277**, 55-64 (2020).

[10] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).

[11] N. Van Doremalen, T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J. L. Harcourt, N. J. Thornburg, S. I. Gerber, *et al.*, Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. *New England Journal of Medicine* **382**, 1564–1567 (2020).

[12] S. Moore, E. M. Hill, M. J. Tildesley, L. Dyson, M. J. Keeling, Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases* (2021).

[13] S. Bauer, S. Contreras, J. Dehning, M. Linden, E. Iftekhar, S. B. Mohr, Á. Olivera-Nappa, V. Priesemann, Relaxing restrictions at the pace of vaccination increases freedom and guards against further COVID-19 waves in Europe. *arXiv preprint arXiv:2103.06228* (2021).

[14] S. Contreras, V. Priesemann, Risking further COVID-19 waves despite vaccination. *The Lancet Infectious Diseases* (2021).

[15] S. Contreras, J. Dehning, M. Loidolt, J. Zierenberg, F. P. Spitzner, J. H. Urrea-Quintero, S. B. Mohr, M. Wilczek, M. Wibral, V. Priesemann, The challenges of containing SARS-CoV-2 via test-trace-and-isolate. *Nature communications* **12**, 1–13 (2021).

[16] M. Linden, S. B. Mohr, J. Dehning, J. Mohring, M. Meyer-Hermann, I. Pigeot, A. Schöbel, V. Priesemann, Case numbers beyond contact tracing capacity are endangering the containment of COVID-19. *Dtsch Arztebl International* **117**, 790-791 (2020).

[17] J. Dehning, J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, V. Priesemann, Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* (2020).

[18] H. W. Hethcote, The Mathematics of Infectious Diseases. *SIAM Review* **42**, 599–653 (2000).

[19] S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, M. H. Wang, Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases* **92**, 214 - 217 (2020).

[20] Y. Alimohamadi, M. Taghdir, M. Sepandi, I. M. A. Wirawan, P. P. Januraga, H.-J. Kim, H.-S. Hwang, Y.-H. Choi, H.-Y. Song, J.-S. Park, Others, Estimate of the Basic Reproduction Number for COVID-19: A Systematic Review and Meta-analysis. *J Prev Med Public Health* **53**, 151–157 (2020).

[21] J. P. La Salle, *The stability of dynamical systems* (SIAM, 1976).

[22] J. Wilting, V. Priesemann, Inferring collective dynamical states from widely unobserved systems. *Nature Communications* **9**, 2325 (2018).

[23] M. J. Keeling, G. Guyver-Fletcher, A. Holmes, L. Dyson, M. J. Tildesley, E. M. Hill, G. F. Medley, Precautionary breaks: planned, limited duration circuit breaks to control the prevalence of COVID-19. *medRxiv* p. 2020.10.13.20211813 (2020).

[24] E. Mahase, COVID-19: Experts recommend two week circuit break to reduce transmission in the UK. *BMJ* **371** (2020).

[25] Z. Kmietowicz, COVID-19: "There is no alternative," says Johnson, announcing new restrictions for England. *BMJ* **371** (2020).

[26] J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, *et al.*, Inferring the effectiveness of government interventions against covid-19. *Science* **371** (2021).

[27] Y. Li, H. Campbell, D. Kulkarni, A. Harpur, M. Nundy, X. Wang, H. Nair, The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *The Lancet Infectious Diseases* (2020).

[28] E. Holt, Slovakia to test all adults for SARS-CoV-2. *The Lancet* **396**, 1386–1387 (2020).

[29] M. J. Mina, R. Parker, D. B. Larremore, Rethinking Covid-19 Test Sensitivity — A Strategy for Containment. *New England Journal of Medicine* **0**, null (2020).

[30] D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, R. Parker, Test sensitivity is secondary to frequency and turnaround time for covid-19 screening. *Science advances* **7**, eabd5393 (2021).

[31] R. D. Kirkcaldy, B. A. King, J. T. Brooks, COVID-19 and Postinfection Immunity: Limited Evidence, Many Remaining Questions. *JAMA* **323**, 2245-2246 (2020).

[32] M. Thompson, J. Burgess, A. Naleway, *et al.*, Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers — eight u.s. locations. *MMWR Morb Mortal Wkly Rep 2021;70:495–500* (2021).

[33] N. Dagan, N. Barda, E. Kepten, O. Miron, S. Perchik, M. A. Katz, M. A. Hernán, M. Lipsitch, B. Reis, R. D. Balicer, BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *New England Journal of Medicine* **384**, 1412–1423 (2021).

[34] M. Voysey, S. A. C. Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, B. Angus, V. L. Baillie, S. L. Barnabas, Q. E. Bhorat, S. Bibi, C. Briner, P. Cicconi, A. M. Collins, R. Colin-Jones, C. L. Cutland, T. C. Darton, K. Dheda, C. J. A. Duncan, K. R. W. Emary, K. J. Ewer, L. Fairlie, S. N. Faust, S. Feng, D. M. Ferreira, A. Finn, A. L. Goodman, C. M. Green, C. A. Green, P. T. Heath, C. Hill, H. Hill, I. Hirsch, *et al.*, Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet* **397**, 99-111 (2021).

[35] E. O.-O. Max Roser, Hannah Ritchie, J. Hasell, Coronavirus pandemic (COVID-19). *Our World in Data* (2020). `https://ourworldindata.org/coronavirus`, (Europe, America, and Oceania and Asia).

[36] T. Hale, S. Webster, A. Petherick, T. Phillips, B. Kira, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government. , `https://www.fda.gov/media/136472/download` (2020).

[37] A. J. Kucharski, P. Klepac, A. J. Conlan, S. M. Kissler, M. L. Tang, H. Fry, J. R. Gog, W. J. Edmunds, J. C. Emery, G. Medley, *et al.*, Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases* **20**, 1151–1160 (2020).

[38] C. C. Kerr, D. Mistry, R. M. Stuart, K. Rosenfeld, G. R. Hart, R. C. Núñez, J. A. Cohen, P. Selvaraj, R. G. Abeysuriya, M. Jastrzębski, *et al.*, Controlling covid-19 via test-trace-quarantine. *Nature communications* **12**, 1–12 (2021).

[39] C. Fraser, S. Riley, R. M. Anderson, N. M. Ferguson, Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* **101**, 6146–6151 (2004).

[40] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368** (2020).

[41] D. Lunz, G. Batt, J. Ruess, To quarantine, or not to quarantine: A theoretical framework for disease control via contact tracing. *Epidemics* **34**, 100428 (2021).

[42] S. Sturniolo, W. Waites, T. Colbourn, D. Manheim, J. Panovska-Griffiths, Testing, tracing and isolation in compartmental models. *PLoS computational biology* **17**, e1008633 (2021).

[43] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, F. Sun, Others, Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* (2020).

[44] E. L. Davis, T. C. D. Lucas, A. Borlase, T. M. Pollington, S. Abbott, D. Ayabina, T. Crellen, J. Hellewell, L. Pi, G. F. Medley, Others, An imperfect tool: COVID-19'test & trace' success relies on minimising the impact of false negatives and continuation of physical distancing. *medRxiv* (2020).

[45] J. J. Van Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, *et al.*, Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour* pp. 1–12 (2020).

[46] S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, T. Wu, The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).

[47] M. Sharma, S. Mindermann, C. Rogers-Smith, G. Leech, B. Snodin, J. Ahuja, J. B. Sandbrink, J. T. Monrad, G. Altman, G. Dhaliwal, L. Finnveden, A. J. Norman, S. B. Oehm, J. F. Sandkühler, T. Mellan, J. Kulveit, L. Chindelevitch, S. Flaxman, Y. Gal, S. Mishra, J. M. Brauner, S. Bhatt, Understanding the effectiveness of government interventions in Europe's second wave of COVID-19. *medRxiv* p. 2021.03.25.21254330 (2021).

[48] D. K. Chu, E. A. Akl, S. Duda, K. Solo, S. Yaacoub, H. J. Schünemann, D. K. Chu, E. A. Akl, A. El-harakeh, A. Bognanni, T. Lotfi, M. Loeb, A. Hajizadeh, A. Bak, A. Izcovich, C. A. Cuello-Garcia, C. Chen, D. J. Harris, E. Borowiack, F. Chamseddine, F. Schünemann, G. P. Morgano, G. E. U. M. Schünemann, G. Chen, H. Zhao, I. Neumann, J. Chan, J. Khabsa, L. Hneiny, L. Harrison, M. Smith, N. Rizk, P. G. Rossi, P. AbiHanna, R. El-khoury, R. Stalteri, T. Baldeh, T. Piggott, Y. Zhang, Z. Saad, A. Khamis, M. Reinap, S. Duda, K. Solo, S. Yaacoub, H. J. Schünemann, Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* **395**, 1973–1987 (2020).

[49] J. Howard, A. Huang, Z. Li, Z. Tufekci, V. Zdimal, H.-M. van der Westhuizen, A. von Delft, A. Price, L. Fridman, L.-H. Tang, *et al.*, An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences* **118** (2021).

[50] J. H. McDermott, W. G. Newman, Refusal of viral testing during the SARS-CoV-2 pandemic. *Clinical Medicine* (2020).

[51] C.-C. Lai, Y. H. Liu, C.-Y. Wang, Y.-H. Wang, S.-C. Hsueh, M.-Y. Yen, W.-C. Ko, P.-R. Hsueh, Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): facts and myths. *Journal of Microbiology, Immunology and Infection* (2020).

[52] E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C. Del Vecchio, L. Rossi, R. Manganelli, A. Loregian, N. Navarin, Others, Suppression of COVID-19 outbreak in the municipality of Vo, Italy. *Nature* (2020).

[53] O. Byambasuren, M. Cardona, K. Bell, J. Clark, M.-L. McLaws, P. Glasziou, Estimating the extent of asymptomatic covid-19 and its potential for community transmission: systematic review and meta-analysis. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada* **5**, 223–234 (2020).

[54] M. Cevik, K. Kuppalli, J. Kindrachuk, M. Peiris, Virology, transmission, and pathogenesis of SARS-CoV-2. *BMJ* **371** (2020).

[55] M. Pollán, B. Pérez-Gómez, R. Pastor-Barriuso, J. Oteo, M. A. Hernán, M. Pérez-Olmeda, J. L. Sanmartín, A. Fernández-García, I. Cruz, N. F. de Larrea, *et al.*, Prevalence of sars-cov-2 in spain (ene-covid): a nationwide, population-based seroepidemiological study. *The Lancet* **396**, 535–544 (2020).

[56] J. A. Firth, J. Hellewell, P. Klepac, S. M. Kissler, A. J. Kucharski, L. G. Spurgin, C. C.-. Working group, Others, Combining fine-scale social contact data with epidemic modelling reveals interactions between contact tracing, quarantine, testing and physical distancing for controlling COVID-19. *medRxiv* (2020).

[57] S. Kojaku, L. Hébert-Dufresne, E. Mones, S. Lehmann, Y.-Y. Ahn, The effectiveness of backward contact tracing in networks. *Nature Physics* **17**, 652–658 (2021).

[58] F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Pérez Marc, E. D. Moreira, C. Zerbini, *et al.*, Safety and efficacy of the bnt162b2 mrna covid-19 vaccine. *New England Journal of Medicine* **383**, 2603–2615 (2020).

[59] M. an der Heiden, O. Hamouda, Schätzung der aktuellen Entwicklung der SARS-CoV-2- Epidemie in Deutschland – Nowcasting. *Epidemiologisches Bulletin* **2020**, 10–15 (2020).

[60] L. F. Shampine, S. Thompson, Solving ddes in matlab. *Applied Numerical Mathematics* **37**, 441–458 (2001).

[61] E. Jarlebring, Some numerical methods to compute the eigenvalues of a time-delay system using MATLAB. *The delay e-letter* **2**, 155 (2008).

[62] L. N. Trefethen, *Spectral methods in MATLAB* (SIAM, 2000).

[63] X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, Others, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* pp. 1–4 (2020).

[64] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R. L. Hesketh, L. Yang, Others, Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology* p. 200370 (2020).

[65] Y. M. Bar-On, A. Flamholz, R. Phillips, R. Milo, Science forum: SARS-CoV-2 (COVID-19) by the numbers. *Elife* **9**, e57309 (2020).

[66] J. Dehning, Priesemann-Group/covid19_metastability. *Zenodo* (2021).

[67] A. F. Siegenfeld, Y. Bar-Yam, The impact of travel and timing in eliminating COVID-19. *Communications Physics* **3**, 204 (2020).

[68] P. Bittihn, R. Golestanian, Containment strategy for an epidemic based on fluctuations in the SIR model. *arXiv:2003.08784 [physics, q-bio]* (2020). ArXiv: 2003.08784.

[69] J. Salvatier, T. V. Wiecki, C. Fonnesbeck, Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2**, e55 (2016).

Author Contributions

Conceptualization: S.C., J.D., V.P.
Methodology: S.C., J.D., V.P.
Software: S.C., S.B.M.
Validation: S.C., J.D., S.B.M., S.B., V.P.
Formal analysis: S.C., J.D., S.B.M., P.S., S.B., V.P.
Investigation: S.C., J.D., S.B.M., P.S., S.B.

Writing - Original Draft: S.C., J.D., S.B.M., P.S., V.P.
Writing - Review & Editing: S.C., J.D., S.B.M., P.S., S.B., V.P.
Visualization: S.C., P.S., S.B.M.
Supervision: V.P.

Competing Interests

V.P. is currently active in various groups to advise the government.

Acknowledgements

We thank Melanie Brinkmann and Álvaro Olivera-Nappa for their helpful comments and encouraging feedback. We thank M. Loidolt, Michael Wibral, Johannes Zierenberg, Joel Wagner, and Anamaria Sanchez for carefully reading, commenting and improving the manuscript. We thank the Priesemann group for exciting discussions and for their valuable input. **Funding:** All authors received support from the Max-Planck-Society (Max-Planck-Gesellschaft MPRG-Priesemann). S.C. acknowledges funding by the Centre for Biotechnology and Bioengineering - CeBiB (PIA project FB0001, ANID, Chile). J.D. and P.S. acknowledge funding by SMARTSTART, the joint training program in computational neuroscience by the VolkswagenStiftung and the Bernstein Network. S.B.M and S.B were financially supported by the German Federal Ministry of Education and Research (BMBF) as part of the Network University Medicine (NUM), project egePan, funding code: 01KX2021.

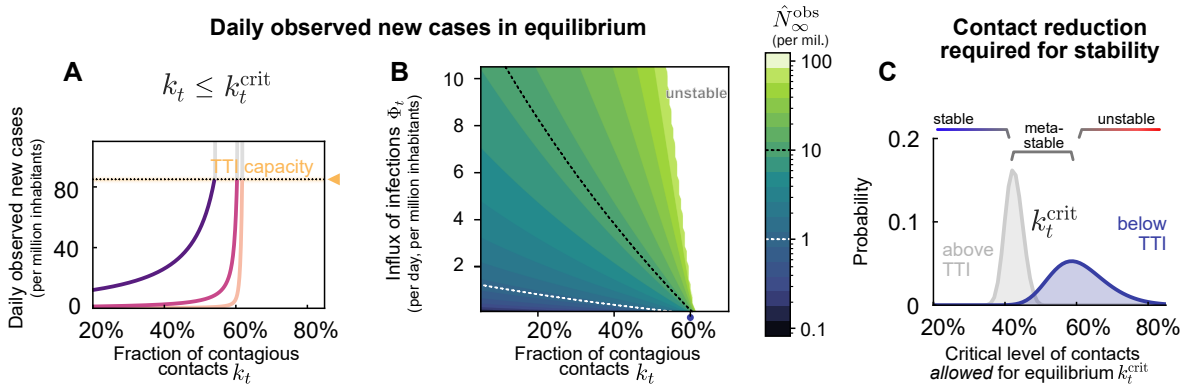
Data and Materials Availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw data and preliminary visualizations of Fig. 6, S4, and S5 were obtained from (*35*). Analysis code available in `https://github.com/Priesemann-Group/covid19_metastability` and permanently stored in (*66*). Additionally, an interactive platform to simulate scenarios different from those presented here `http://covid19-metastability.ds.mpg.de/`.

List of Supplementary Materials

- **Supplementary Note S1.1:** Strategies to face COVID-19 differ among countries.
- **Supplementary Note S1.2:** Linear stability analysis and uncertainty propagation.
- **Supplementary Note S1.3:** On the contact reduction required for achieving early population immunity.
- **Supplementary Note S1.4:** Calculating the increase in the level of contacts allowed with increased immunity.
- **Supplementary Note S1.5:** Inferring the reproduction number of COVID-19 from Jun to Oct. 2020.
- **Supplementary Note S1.6:** On the incorporation of random testing in the TTI scheme.
- **Supplementary Note S1.7:** More analytical insights into the TTI-based metastable regime.
- **Supplementary Note S1.8:** Exploring the effect of more compartments for the exposed individuals.

S1 Supplementary Notes of "Low case numbers enable long-term stable pandemic control without lockdowns"



Supplementary Figure S1: **In the stable and metastable regimes, daily new cases approach an equilibrium value $\hat{N}_\infty^{\text{obs}}$ that depends on the level of contacts k_t and an external influx of new cases Φ_t. A:** The equilibrium value $\hat{N}_\infty^{\text{obs}}$ increases steeply with higher k_t and then destabilizes when surpassing a critical value k_t^{crit}. Thus, contact reduction has to be sufficiently strong to keep the case numbers within TTI capacity. A certain degree of external influx Φ_t can be compensated, but in general, Φ_t can put stability at risk. $\Phi_t = 1$ one daily new case per million is our default parameter for the influx. **B:** The $\hat{N}_\infty^{\text{obs}}$ is below 10 for a large fraction of combinations of k_t and Φ_t, thus well below the capacity limit of TTI. **C:** The critical value k_t^{crit}, which marks the transition between stable and (meta)stable spread; thus k_t^{crit} displays the maximal level of contacts that is allowed while controlling the spread (and stabilize case numbers). If case numbers are below the TTI capacity limit, a contact level of at most $k_t^{\text{crit}} = 61\,\%$ is allowed for stabilization (blue). If case numbers, however, are above the TTI limit, stronger contact reduction is necessary for stabilization, thus allowing a lower level of contacts ($k_t^{\text{crit}} = 42\,\%$, gray). Confidence intervals originate from error propagation of the uncertainty of the underlying model parameters. All model parameters are listed in Supplementary Table S1 and the full uncertainty analysis is in Fig. S2.

S1.1 Strategies to face COVID-19 differ among countries.

Several South- and East-Asian countries and Australia, and New Zealand have achieved very low case numbers and even local eradication. These countries reached very low values below one daily new case per million (median 0.5, Fig. 6 C). If local eradication is successful, these countries can profit from the absorbing state of zero SARS-CoV-2 infections, i.e. after local eradication, new infection chains are only started if a virus is *de novo* carried into the country (*67, 68*). However, the local eradication is constantly put at risk by the undetected influx of new viruses from abroad, requiring rigorous quarantine for international travel, and – once the spread got out of control – decisive action to completely stop all infection chains. However, the more countries adhere to this strategy successfully, the closer one may get to global eradication.

In many European countries over the summer, case numbers were relatively low, typically around ten daily new cases per million (Fig. 6 B). During that time, contacts were only mildly or moderately restricted, and containment was complemented by hygiene, masks, and other preventive measures. However, in summer and autumn, most European countries developed a second wave Fig. S5. The causes are undoubtedly diverse, from increasing contact rates to seasonal effects and travel-related influx. Seasonal effects alone cannot explain the second wave, as neighboring countries like Portugal versus Spain or Finland versus Sweden show remarkably different dynamics (see (*35*)). Hence, it seems to be possible to maintain an equilibrium at relatively low case numbers. However, that equilibrium is fragile at high case numbers, and novel waves can emerge at any time.

Sustained high levels of case numbers have been observed in several countries such that TTI probably could not be performed effectively. Around 130 daily new cases per million have been observed, e.g. in many (but not all) American countries (median 129.54, Fig. 6 A). It shows that high levels of daily new infections

can be maintained in principle. However the stringency of interventions is similar or higher compared to other countries (see (*36*)), and even with these high numbers it will probably take about $200\,000/150 = 1333$ to $700\,000/150 = 4666$ days, thus several years, until 20 to 70 % of the population is infected and population immunity reached – assuming the duration of individual immunity is long enough. That high level of new infections leads to a considerable death toll, as currently about 1.5 % of the infected individuals would die (depending on age structure (*2, 16*)). Moreover, containment measures like quarantine become unsustainable because, if implemented, each one of the 200 daily new infected cases would require the quarantine of 5-50 people (their high-risk contacts) for about ten days, causing 1 - 10 % of the population being in quarantine at any given day. Therefore the alleged economic and social benefits of such a strategy (*3, 7*) may be questionable.

S1.2 Linear stability analysis and uncertainty propagation

For analyzing the stability of the governing differential equations, namely, whether an outbreak could be controlled, we studied the linear stability of the system. Moreover, we consider that, within the time-frame considered for stability purposes, the fraction $\frac{S}{M}$ would remain somewhat constant, we consider the linearized version of equations (3)– (7), defining a system of delay differential equations for the variables $x(t) = \left[E^Q(t);\, E^H(t);\, I^Q(t);\, I^H(t);\, I^{H,s}(t)\right]$. We define matrices A and B as:

$$A = \begin{pmatrix} -\rho & 0 & \nu\gamma R_0 & 0 & 0 \\ 0 & -\rho & \epsilon\gamma R_0 & \gamma k_t R_0 & 0 \\ \rho & 0 & -\gamma & \lambda_r & \lambda_s \\ 0 & \rho & 0 & -\gamma-\lambda_r & -\lambda_s \\ 0 & (1-\xi)\,\rho & 0 & 0 & -\gamma-\lambda_r-\lambda_s \end{pmatrix} \tag{1}$$

$$B = \begin{pmatrix} 0 & 0 & 0 & \lambda_r^{\text{eff}}\chi_\tau & \lambda_s^{\text{eff}}\chi_\tau \\ 0 & 0 & 0 & -\lambda_r^{\text{eff}}\chi_\tau & -\lambda_s^{\text{eff}}\chi_\tau \\ 0 & 0 & 0 & \lambda_r^{\text{eff}}\left(\xi\chi_r + (1-\xi)\,\chi_{s,r}\right) & \lambda_s^{\text{eff}}\left(\xi\chi_r + (1-\xi)\,\chi_{s,r}\right) \\ 0 & 0 & 0 & -\lambda_r^{\text{eff}}\left(\xi\chi_r + (1-\xi)\,\chi_{s,r}\right) & -\lambda_s^{\text{eff}}\left(\xi\chi_r + (1-\xi)\,\chi_{s,r}\right) \\ 0 & 0 & 0 & -\lambda_r^{\text{eff}}\,(1-\xi)\,\chi_{s,r} & -\lambda_s^{\text{eff}}\,(1-\xi)\,\chi_{s,r} \end{pmatrix} \eta k_t R_0, \tag{2}$$

where

$$\lambda_r^{\text{eff}} = \frac{\gamma\lambda_r}{\lambda_r+\gamma}, \qquad \lambda_s^{\text{eff}} = \gamma\left(\frac{\lambda_s+\lambda_r}{\gamma+\lambda_s+\lambda_r} - \frac{\lambda_r}{\lambda_r+\gamma}\right). \tag{3}$$

The equations governing the dynamics for vector $x(t)$ are then presented in their matrix form:

$$x'(t) = Ax(t) + Bx(t-\tau). \tag{4}$$

We determine the maximum –critical– level of contacts k_t^{crit}, for which exponential solutions would be asymptotically stable. Eigenvalues were determined by systematically solving the nonlinear eigenvalues problem for stability (*61*), where the solution operation was approximated with a Chebyshev differentiation matrix (*62*). Eigenvalues, in this sense, would be solutions of the scalar equation

$$\det\left(-sI + A + e^{-s\tau}B\right) = 0 \tag{5}$$

Noting that A and B explicitly depend on the model parameters, we numerically explore which combinations would result in stable, metastable, or unstable case numbers. Concretely, we studied the maximum –critical– level of contacts allowed for stability k_t^{crit} in two scenarios; i) low case numbers, and TTI fully operative (both testing and contact tracing), and ii) high case numbers, above the TTI limit, where testing would be inefficient and solely symptom driven ($\lambda_s = \lambda_s'$, $\eta = \lambda_r = \lambda_r' = 0$).

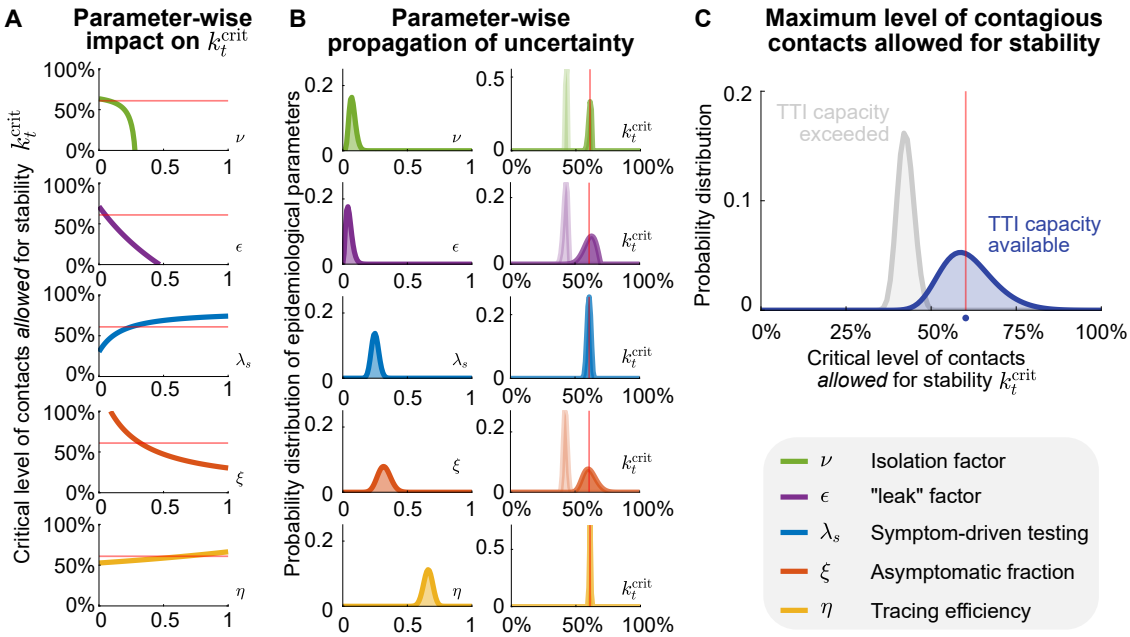
To explore the uni-variate impact different signature parameters have on k_t^{crit}, we studied the zeros of (5) as a function of k_t using the `@fzero` MATLAB function (Fig. S2A). Using the same routines and a random sampling procedure, we propagate uncertainties in the values of these parameters, uni-variate (Fig. S2B), and multivariate (Fig. S2C).

Supplementary Table S1: Parameter uncertainty propagation. α and β are the shape parameters of the beta distribution.

Parameter	Meaning	Median	95% CI	α	β	Dist.	Units
ξ	Asymptomatic ratio	0.32	0.23–0.42	27.5	27.8	beta	–
λ_s	Symptom-driven test rate	0.25	0.20–0.31	56	168	beta	days^{-1}
ν	Registered contacts (quarantined)	0.07	0.03–0.13	8.25	101.8	beta	–
η	Tracing efficiency	0.66	0.59–0.73	117.9	60.7	beta	–
ϵ	Lost contacts (quarantined)	0.05	0.01–0.11	3.8	71.25	beta	–
$k_t^{\text{crit}}\big\vert_{\text{TTI}}$	Maximum –critical– level of contacts allowed for stability (with TTI)	61 %	47–76 %	–	–	–	–
$k_t^{\text{crit}}\big\vert_{\text{no TTI}}$	Maximum –critical– level of contacts allowed for stability (without TTI)	42 %	38–47 %	–	–	–	–

Supplementary Table S2: Linearly-derived correspondence between contact reduction and the observed reproduction number

k_t	$\hat{R}_t^{\text{obs}}$ (with TTI)	$\hat{R}_t^{\text{obs}}$ (without TTI)
0.8	1.08	1.26
0.6	0.99	1.13
0.4	0.91	0.98
0.25	0.86	0.88



Supplementary Figure S2: **Propagation of TTI-parameter uncertainties to the critical level of contacts allowed for stability.** As the different parameters involved in our model play different roles, the way their variability propagates to k_t^{crit} differs, even when their variability profiles look similar. **A:** Impact of single-parameter variation on the critical (maximal, allowed) level of contagious contacts k_t^{crit}. To evaluate the monotony (direction) of their impact on k_t^{crit}, we scan their entire definition range, ignoring the practical feasibility of achieving such values. The solid red line shows the default critical level of contacts allowed for stability. **B:** Univariate uncertainties of TTI parameters modeled by beta distributions centered on their default value, and the resulting distribution of critical fraction of contagious contacts k_t^{crit} (right column). Results are shown assuming testing only (light colors) or testing and tracing (dark colors). Solid red lines represent the default value of k_t^{crit} in the regime of available TTI capacity. **C:** Distribution of the critical fraction of contacts arising from multivariate uncertainty propagation given by the joint of the distributions shown in (A) for testing only (light colors) or testing and tracing (dark colors). Solid red lines represent the default value of k_t^{crit} in the regime of available TTI capacity. Results show averages of 100 000 realizations.

S1.3 On the contact reduction required for achieving early population immunity.

In Supplementary Section S1.2, we derived a methodology for obtaining the minimal, critical contact reduction k_t^{crit} for which the linear system is asymptotically stable. Such values, however, assume a fully susceptible population, as we ignore the scaling factor $\frac{S}{M}$.

The population immunity threshold ϱ represents the fraction of the population that needs to be immunized for controlling the spread of an infectious disease. It can be expressed in terms of the effective reproduction number R_t:

$$\varrho = 1 - \frac{1}{R_t}. \tag{6}$$

In the context of our model, R_t can be expressed in terms of the fraction of contagious contacts k_t and the basic reproduction number R_0; $R_t = k_t R_0$. However, in further stages of an ongoing outbreak, the fraction of people no longer susceptible would affect the spread. Thus we include also the scaling factor $\frac{S}{M}$:

$$R_t = k_t \frac{S}{M} R_0. \tag{7}$$

Combining both equation (6) and (7), we can express ϱ as

$$\varrho = 1 - \frac{1}{k_t\,(1-f)\,R_0}, \tag{8}$$

assuming a quasi-stationary dynamics for S, and defining $f = 1 - \frac{S}{M}$. Suppose we study the case in which no major behavioral changes take place. Thus the population immunity threshold ϱ would remain the same. On the other hand, because of the sole fact of having a progressively increasing immunization among the population (because of vaccination or post-infection immunity), the maximal allowed level of contacts k_t^{crit} will increase. Assuming critical conditions, we use $k_t = k_t^{\text{crit}}$ in equation (8):

$$\varrho = 1 - \frac{1}{k_t^{\text{crit}}\,(1-f)\,R_0}, \tag{9}$$

As we assumed that no behavioral change is taking place, we obtain the population immunity threshold by only evaluating equation (9) at $f = 0$.

$$\varrho = 1 - \frac{1}{k^{\text{crit},0} R_0}, \tag{10}$$

where $k^{\text{crit},0}$ represents the critical level of contacts allowed in a fully susceptible population, and can be calculated directly from the linear stability analysis described in S1.2. Subtracting (9) and (10) we obtain an expression for $k_t^{\text{crit}}(f)$

$$k_t^{\text{crit}}(f) = \frac{k^{\text{crit},0}}{1-f}. \tag{11}$$

Finally, this expression only depends on the remaining susceptible population $(1-f)$, and on the critical level of contacts when $f = 0$. Note that this equation can return values of $k_t^{\text{crit}}(f)$ larger than one, should f be *close enough* to one. Following our interpretation of $k_t^{\text{crit}}(f)$, that would mean reaching the population immunity level, as individuals would be allowed to have even higher levels of contacts than those they had before COVID-19.

S1.4 Calculating the increase in the level of contacts allowed with increased immunity

In the previous section, we demonstrated that with increased immunity (acquired by vaccination or post-infection), the maximum –critical– level of contacts k_t^{crit} allowed for stability would increase. We also discussed that, as a first-order approach, the level at which the system is stabilized (namely, $\hat{N}_\infty^{\text{obs}}$) is determined by the influx Φ_t and the distance between the current level of contacts k_t and k_t^{crit} according to the formula

$$\hat{N}_\infty^{\text{obs}} = \frac{\Phi_t}{k_t^{\text{crit}} - k_t}. \tag{12}$$

We can re-write equation (12) as:

$$k_t = \frac{k^{\text{crit},0}}{1-f} - \frac{\Phi_t}{\hat{N}_\infty^{\text{obs}}}, \tag{13}$$

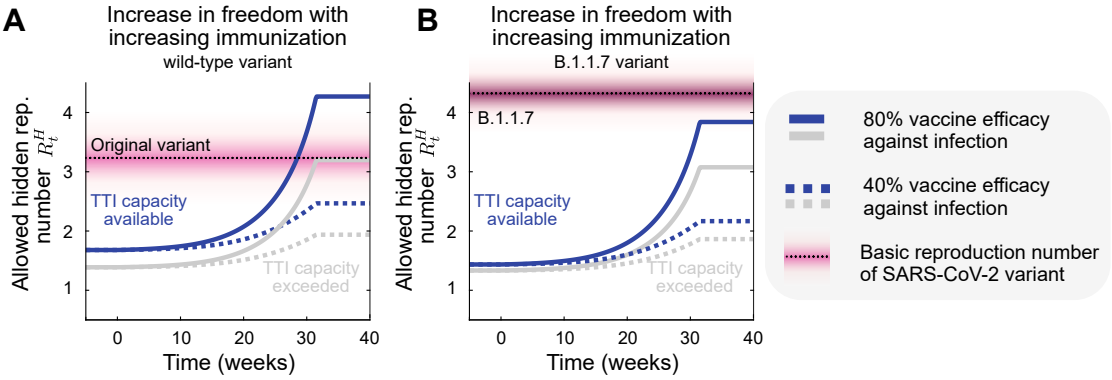
where k_t represents the *allowed* level of contagious contacts so that the distance to the singularity (reached if $k_t = k_t^{\text{crit}}$) is kept constant. According to equation (13), increasing levels of immunity will lead to higher critical level of contacts, thus allowing individuals to steadily increase the level of contagious contacts k_t while still keeping the same equilibrium value for case numbers.

In order to connect this quantity to the predominant circulating variants, we recall the definition of the hidden reproduction number R_t^H (which is slightly different from the effective reproduction number R_t

defined in equation (7)). As described in the main text and in (*15*), R_t^H accounts for the number of offspring infections generated by individuals unaware of being infectious, in a fully susceptible population. In that way, we can express it in terms of k_t; $R_t^H = k_t R_0$. Thus, we can rewrite equation (13) in terms of the allowed hidden reproduction number R_t^H, just by multiplying it by R_0.

$$R_t^H = \frac{k_t^{\text{crit},0} R_0}{1-f} - \frac{\Phi_t R_0}{\hat{N}_\infty^{\text{obs}}}, \tag{14}$$

We note a slight coupling between the hidden reproduction number R_t^H and the dominant variant of SARS-CoV-2 (through the base reproduction number), so scenarios need to be analyzed separately. Using the vaccination progress described in Fig. 5 in the main text, our results are summarized in Fig. S3, where we assume that the TTI and hospital strategies (resp. blue and gray curves) stabilize at 10 and 250 cases per million people, respectively. The influx term that reduces the allowed hidden reproduction number R_t^H highlights the need of combining NPIs aiming to lower case numbers to the TTI regime (where $k_t^{\text{crit},0}$ is higher), especially when large fractions of the population remain susceptible (cf. Fig. S3A). However, in the presence of highly contagious variants, NPIs must remain in place, as even after vaccine rollout the allowed hidden reproduction number could not surpass the base reproduction number of the variant. Even though reducing the influx could also help, its effect is little compared with the effect of immunization (cf. Fig. S3B).



Supplementary Figure S3: **As growing immunity increases the maximal allowed number of contacts, freedom will proportionally increase.** Scenarios considering the original, wild-type variant of SARS-CoV-2 ($R_0 = 3.3$) show that vaccination together with a strategy aiming at low case numbers can make the allowed hidden reproduction number R_t^H cross the variant's R_0, i.e., the population immunity threshold (**A**). However, considering the higher reproduction number of the B.1.1.7, restrictions are not allowed to be lifted fully to keep case numbers under control, even if within TTI capacity (**B**). The vaccination progress here was adapted from Figure 5 from the main text.

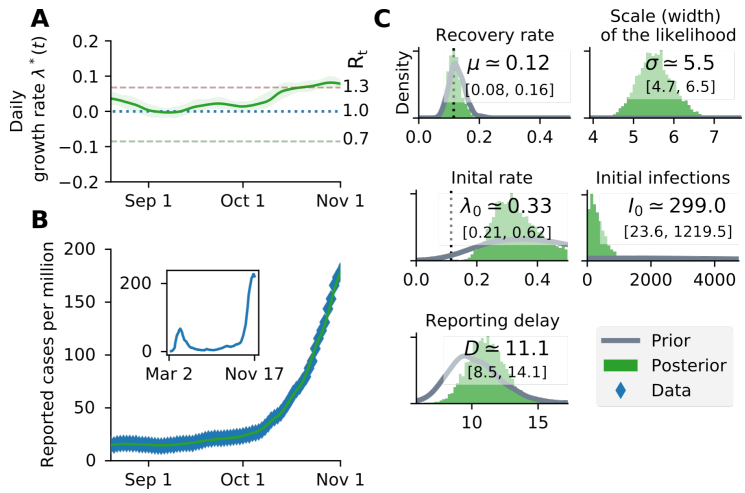
S1.5 Inferring the reproduction number of COVID-19 from Jun to Oct. 2020

We use the Bayesian inference framework on an SIR model presented in our previous work (*17*) to infer the daily growth rate $\lambda^*(t)$. An SIR compartmental model with weekly change points is used whose main epidemiological parameters are inferred using the PyMC3 package (*69*). A weekly modulation was applied to take the weekly reporting structure and weekend delays into account. After inference, a rolling average was performed on the daily case numbers for comparability and clarity. The observed reproduction number $\hat{R}_t^{\text{obs}}$ can be expressed depending on the effective daily growth rate:

$$\hat{R}_t^{\text{obs}} = (\lambda^*(t) + 1)^4 \tag{15}$$

This short analysis was performed for Germany (Fig. S4) and other European countries (Fig. S5), which showed the same metastable behavior, while their case numbers were below a threshold of around 50 cases

per day (per million). A transition into the unstable regime can be seen once case numbers surpass this threshold.



Supplementary Figure S4: **A substantial increase in Germany's effective growth rate occurred during October, suggesting that regional TTI capacity limits were exceeded.** (**A**) Before October, the reproduction number was slightly above one (corresponding to a daily growth rate λ^* slightly above zero). (**B**) Observed case numbers were stabilized below 20 daily new cases per million (but still slowly growing) until a transition into the unstable regime took place over ~ 4 weeks in October. The time range is adjusted to focus on this tipping point. The inset shows case numbers for the full available time range. (**C**) Main central epidemiological parameters with the prior and posterior distribution.

S1.6 On the incorporation of random testing in the TTI scheme

Testing can also be done randomly or randomly combined with a contact tracing strategy. Even though the number of tests required for such purposes would be enormous, the development of fast, cheap, and reliable tests offers an exciting alternative to consider.

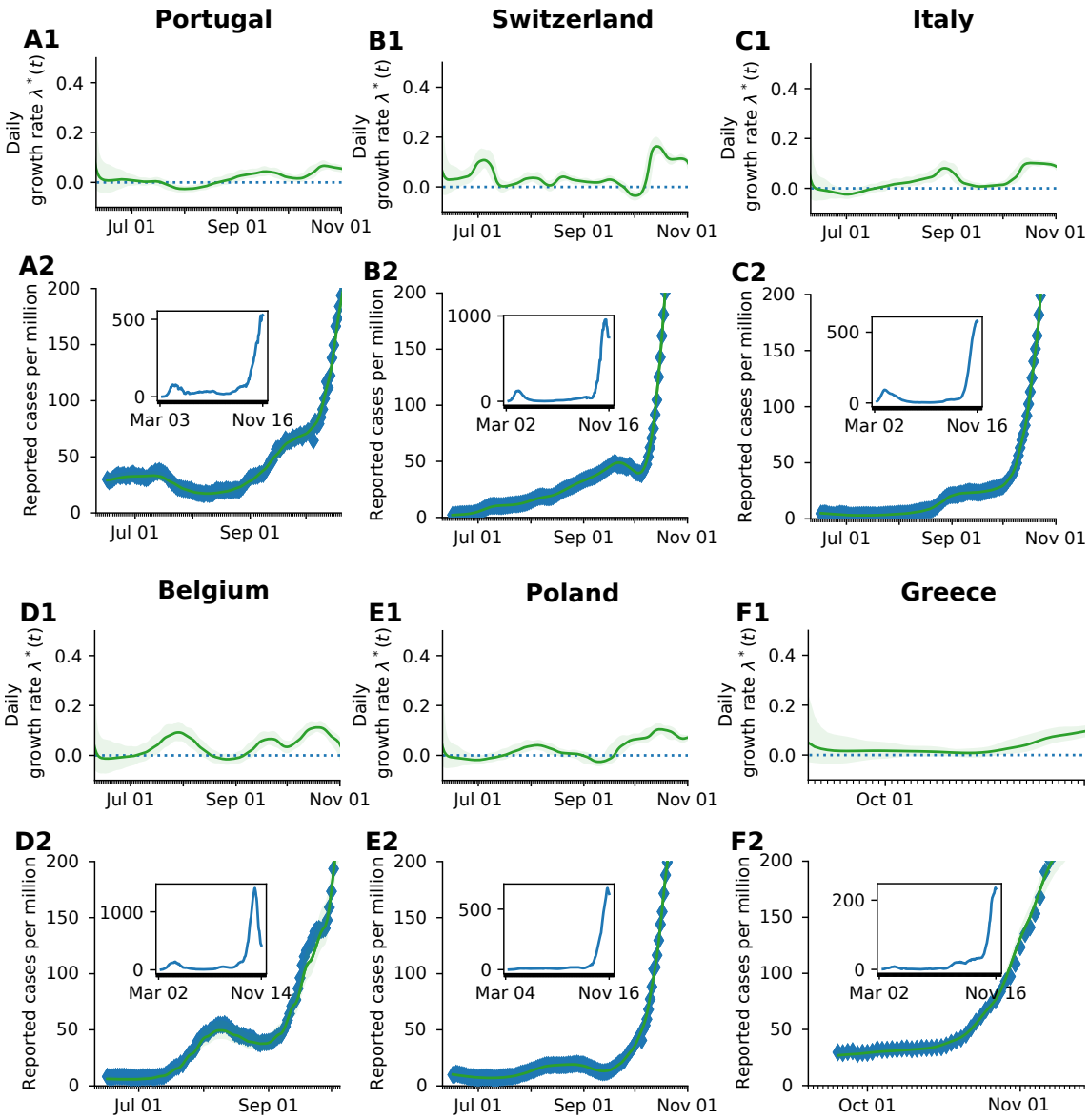
In this supplementary note, we derive the equations presented in the methods but also including random testing. As therein described, random testing is assumed to occur at a constant rate λ_r, which, for default parameters, reflect the number of tests per day per million people. In that way, it seems reasonable to consider unfeasible testing rates surpassing $\lambda_r^{\max} = 0.1$, as it would mean that a 10 % of the population is tested every day.

S1.6.1 Number of cases observable through testing N^{test}

When random testing is included in the scheme, the solution of equation (19) for the symptomatic and asymptomatic infections -hidden- would be given by equations (16) and (17).

$$\chi_{s,r} = \begin{cases} \rho\tau \exp\left(-\rho\tau\right), & \text{if } \rho \approx \lambda_s + \lambda_r \\ \dfrac{\rho}{(\lambda_s + \lambda_r) - \rho}\left(\exp\left(-\rho\tau\right) - \exp\left(-\tau\left(\lambda_s + \lambda_r\right)\right)\right), & \text{else.} \end{cases} \tag{16}$$

$$\chi_r = \begin{cases} \rho\tau \exp\left(-\rho\tau\right), & \text{if } \rho \approx \lambda_r \\ \dfrac{\rho}{\lambda_r - \rho}\left(\exp\left(-\rho\tau\right) - \exp\left(-\tau\lambda_r\right)\right), & \text{else.} \end{cases} \tag{17}$$



Supplementary Figure S5: **Comparison of the reproduction number and reported cases as the second wave emerges in different European countries.** For each country, parameters of an SIR model, were fitted to the reported data of the Our World in Data repository (*35*), following the procedure presented in (*17*). (**Panels X1**) The time-dependent effective growth rate stays between -0.1 and 0.1 and rises before the tipping point. This corresponds to an effective reproduction number between 0.7 and 1.3, which matches our preliminary assumptions. The time range is adjusted to focus on the tipping point. (**Panels X2**) After a (meta-)stable regime in summer, all selected countries show a rise in case numbers and a tipping point at around 50 new cases per day per million. The spread self-accelerates, and the cases increase significantly. (**Insets**) Case numbers for the full available time range.

If both symptom-based and random testing take place simultaneously, the number of discovered infections is given by

$$N^{\text{test}} = \lambda_r I^H + \lambda_s I^{H,s} \tag{18}$$

Further, assuming that after reaching $N^{\text{test}}_{\text{max}}$, the testing rates at the overhead pool-sizes would decrease to λ'_s and λ'_r, respectively, for symptom-driven and random testing. We further assume that testing resources

would be exclusively allocated to sustain the symptom-driven testing in our default scenario. The overall testing term N^{test} would be given by:

$$\begin{aligned} N^{\text{test}} = \lambda_r \min\left(I^H, I^H_{\max}\right) + \lambda'_r \max\left(0, I^H - I^H_{\max}\right) \\ + \lambda_s \min\left(I^{H,s}, I^{H,s}_{\max}\right) + \lambda'_s \max\left(0, I^{H,s} - I^{H,s}_{\max}\right), \end{aligned} \tag{19}$$

where $I^{H,s}_{\max}, I^H_{\max}$ represent the pool sizes of the symptomatic hidden and total hidden pools, respectively, at the TTI limit, i.e. $\lambda_r I^H_{\max} + \lambda_s I^{H,s}_{\max} \overset{!}{=} N^{\text{test}}_{\max}$, reached at time $t = t^*$. Defining $\varphi := \left.\frac{I^{H,s}}{I^H}\right|_{t=t^*}$, we can express such magnitudes in term of the maximum capacity $N^{\text{test}}_{\max}$:

$$I^{H,s}_{\max} = \frac{\varphi N^{\text{test}}_{\max}}{\varphi\lambda_s + \lambda_r} \tag{20}$$

$$I^H_{\max} = \frac{N^{\text{test}}_{\max}}{\varphi\lambda_s + \lambda_r}. \tag{21}$$

The explicit value of φ can be obtained numerically in the integration routine or estimated through the use of the equilibrium values of the differential equations, $\varphi = \frac{I^{H,s}_\infty}{I^H_\infty}$ (as implemented in our code). The expression for the symptomatic hidden pool $I^{H,s}$ in the presence of random testing is slightly different;

$$\begin{aligned} N^{\text{test}}_s = \lambda_r \min\left(I^{H,s}, I^{H,s}_{\max}\right) + \lambda'_r \max\left(0, I^{H,s} - I^{H,s}_{\max}\right) \\ + \lambda_s \min\left(I^{H,s}, I^{H,s}_{\max}\right) + \lambda'_s \max\left(0, I^{H,s} - I^{H,s}_{\max}\right). \end{aligned} \tag{22}$$

S1.6.2 Random testing: number of cases observable through contact-tracing N^{traced}

Because of TTI, infectious individuals move (or are likely to move) from the hidden to the quarantined infectious pool before recovering. Therefore, they spend a comparatively shorter amount of time there and, on average, would not generate the expected amount of offspring infections as some would be prevented. In the absence of TTI, the average time individuals spend in the infectious compartment is $\frac{1}{\gamma}$. In the presence of selective TTI, symptomatic individuals would have a greater chance to be tested (and thereby removed) than the asymptomatic ones. As we explicitly consider compartments for symptomatic and asymptomatic infections, each pool's residence times would be different. Noting that symptomatic individuals can be tested and therefore removed by any of the testing criteria, their residence time would be approximate $\frac{1}{\gamma+\lambda_s+\lambda_r}$. In contrast, the average residence time of asymptomatic individuals would be $\frac{1}{\gamma+\lambda_r}$. Therefore, the fractions of time that symptomatic and asymptomatic individuals stay unnoticed are respectively

$$t_s = \frac{\gamma}{\gamma + \lambda_s + \lambda_r}, \tag{23}$$

and

$$t_r = \frac{\gamma}{\gamma + \lambda_r}. \tag{24}$$

If the daily new cases observed through testing, delayed at the moment of processing, $N^{\text{test}}_{t-\tau}$, are within the tracing capacity of the health authorities, i.e. $N^{\text{test}}_{t-\tau} \leq N^{\text{test}}_{\max}$, then N^{traced} is defined as

$$N^{\text{traced}} = \eta R_{t-\tau}\left(I^H_{t-\tau} t_r \lambda_r + I^{H,s}_{t-\tau}\left(t_s\lambda_s + (t_s - t_r)\lambda_r\right)\right), \tag{25}$$

where $R_{t-\tau}$ represents the effective reproduction number, as defined in equation (7). Otherwise, using the expressions for I^H and $I^{H,s}$ when the TTI capacity is reached derived in the previous section, we can obtain an effective rate

$$\lambda_{\text{eq}} = \frac{\lambda_r \left((1-\varphi)\, t_r + \varphi t_s \right) + \varphi \lambda_s t_s}{\lambda_r + \varphi \lambda_s}. \tag{26}$$

Therefore, the average amount of positive cases identified by contact tracing in the TTI limit is given by

$$N^{\text{traced}} = \eta R_{t-\tau} N^{\text{test}}_{\max} \lambda_{\text{eq}}. \tag{27}$$

To sum up the last equations:

$$N^{\text{traced}} = \begin{cases} \eta R_{t-\tau} (I^H_{t-\tau} t_r \lambda_r + I^{H,s}_{t-\tau} \left(t_s \lambda_s + (t_s - t_r)\, \lambda_r \right)) & \text{if } N^{\text{test}}_{t-\tau} \leq N^{\text{test}}_{\max} \\ \eta R_{t-\tau} N^{\text{test}}_{\max} \lambda_{\text{eq}} & \text{else} \end{cases} \tag{28}$$

S1.7 More analytical insights into the TTI-based metastable regime

One of the crucial aspects of this paper is the description and profit from a metastable regime at low case numbers. Because of the different compartments at interplay in the dynamics, some metastability aspects are easier to present using a simplified model. The critical element behind the case-number dependent metastability is the limited capacity to perform contact tracing of the newly identified infections. We refer to the TTI limit being reached at a prevalence given by $I = I^{\max}$. However, the fact that this metastable equilibrium shows non-zero case numbers arises from a small but non-vanishing influx of externally acquired infections to the system, which we call Φ_t.

Aiming to illustrate the dynamics in a minimalist way, we proceed as follows. We modify a plain SIR model to include hidden and quarantined infections but simplify the TTI scheme to only depend on two parameters: the TTI capacity referred to the hidden incidence $I^{\max}_H$ and the testing rate λ as long as the capacity is not exceeded. If the capacity is exceeded the testing rate drops to zero. Infections in the quarantined and hidden pools spread at rates β_1 and β_2 respectively, with $\beta_1 < \beta_2$, and the recovery rate is given by γ:

$$\frac{dS}{dt} = -\beta_1 I^Q \frac{S}{M} - \beta_2 I^H \frac{S}{M} - \Phi_t \tag{29}$$

$$\frac{dI^Q}{dt} = \beta_1 I^Q \frac{S}{M} - \gamma I^Q + \min(\lambda I^H, \lambda I^H_{\max}) \tag{30}$$

$$\frac{dI^H}{dt} = \beta_2 I^H \frac{S}{M} - \gamma I^H - \min(\lambda I^H, \lambda I^H_{\max}) + \Phi_t. \tag{31}$$

We observe that the metastable regime's key ingredients relate to TTI: The system reacts less intensively to new infections when case numbers are below the TTI capacity, as infection chains are diligently found and removed. On the other hand, when the TTI system is overwhelmed, infections reproduce at the natural, hidden rate. Assuming quasi-stationary dynamics for the susceptible pool (S), we can approximate:

$$\frac{d}{dI^H}\left(\frac{dI^H}{dt}\right) = \begin{cases} \beta_2 \dfrac{S}{M} - \gamma - \lambda, & \text{if } I^H < I^H_{\max}, \\ \\ \beta_2 \dfrac{S}{M} - \gamma, & \text{otherwise.} \end{cases} \tag{32}$$

If I^H exceeds $I^H_{\max}$, there is the possibility for I^H to grow and lead to a more significant number of cases for a broader range of parameters. On the other hand, when studying the linear system's equilibrium,

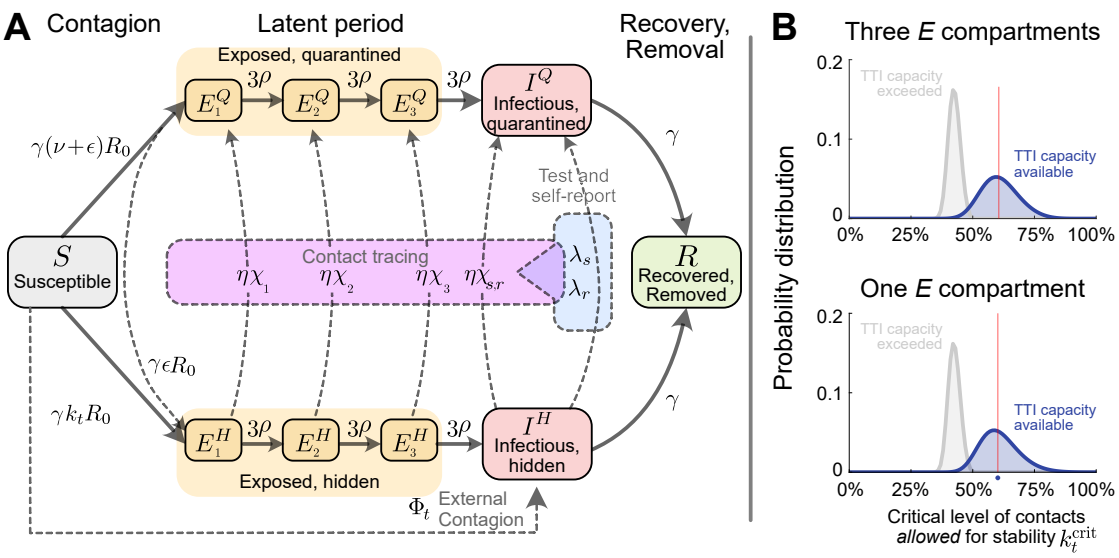
we identify the contributions to the observed number of cases $\hat{N}_\infty^{\text{obs}} = \beta_1 I_\infty^Q + \lambda I_\infty^H$. Assuming stationary conditions for the $\frac{S}{M} \approx 1$ limit, we obtain:

$$\hat{N}_\infty^{\text{obs}} = \frac{\Phi_t}{\lambda + \gamma - \beta_2} \frac{\gamma\lambda}{\gamma - \beta_1}. \tag{33}$$

Using these simple models, we can better understand the parameters and mechanisms that render the system unstable or the equilibrium unfeasible. This better understanding of factors contributing to stabilizing –and respectively destabilizing– the system, helps to guide policies. The transitions and equilibrium levels have more complex analytical expressions in our model but follow the same spirit: to draw the line between the stable and unstable spread.

S1.8 Exploring the effect of more compartments for the exposed individuals

Compartmental models coupled with differential equations implicitly assume first-order kinetics, which translates to exponential emptying or filling the compartments involved in the dynamics. Even though the expected residence time can be set for each compartment, the exponential shape of the residence time distribution differs from the true disease dynamics. This difference can be corrected by including extra compartments, so instead of having an exponential shape, the residence time distribution eventually converges to a delta distribution around the desired value. Aiming to evaluate how this effect might change our estimates for k_t^{crit}, we included three compartments for the exposed fraction of the population (and for hidden and quarantined infections) presented in Fig. S6. As the latent period would be spent in these compartments, the residence time in each one is one-third of the original, so that the new transition rate would be $\rho' = 3\rho$.



Supplementary Figure S6: **Flowchart of the complete model, including three compartments for the latent period. A:** The solid blocks in the diagram represent different SEIR compartments for both hidden and quarantined individuals. Hidden cases are further divided into symptomatic and asymptomatic carriers (not shown). Solid lines represent the natural progression of the infection (contagion, latent period, and recovery). On the other hand, dashed lines account for imperfect quarantine and limited compliance, external factors, and test-trace-and-isolate policies. **B:** Comparison between the distribution of k_t^{crit} for the single-compartment model and the three compartments model.

Following the same formalism presented in the previous sections, we estimate the fraction of individuals infected at time t that would remain in the different compartments by the time of contact tracing.

$$\chi_1 = \exp\left(-3\rho\tau\right), \tag{34}$$

$$\chi_2 = (3\rho)\,\tau \exp\left(-3\rho\tau\right), \tag{35}$$

$$\chi_3 = (3\rho)^2 \frac{\tau^2}{2} \exp\left(-3\rho\tau\right). \tag{36}$$

Consistently, we can demonstrate that the fraction of individuals staying in the infectious compartment in the presence of symptom-based testing is given by

$$\chi_{s,r} = \begin{cases} (3\rho)^3 \dfrac{\tau^3}{3!} \exp\left(-3\rho\tau\right), & \text{if } 3\rho \approx \lambda_s + \lambda_r \\ \dfrac{(3\rho)^3 \exp\left(-3\rho\tau\right)}{2(\lambda_s+\lambda_r-3\rho)} \left(\tau^2 - 2\dfrac{\tau}{\lambda_s+\lambda_r-3\rho} + \dfrac{2}{(\lambda_s+\lambda_r-3\rho)^2}\left(1-\exp\left(-\tau\left(\lambda_s+\lambda_r-3\rho\right)\right)\right)\right), & \text{else.} \end{cases} \tag{37}$$

To analyze the stability of the extended system with three exposed compartments, we proceeded as described in Section S1.2. The linear system with delay representing the dynamics is $x'(t) = Ax(t) + Bx(t-\tau)$, with $x(t) = \left[E_1^Q(t);\, E_2^Q(t);\, E_3^Q(t);\, E_1^H(t);\, E_2^H(t);\, E_3^H(t);\, I^Q(t);\, I^H(t);\, I^{H,s}(t)\right]$, where matrices A and B are given by:

$$A = \begin{pmatrix} -3\rho & 0 & 0 & 0 & 0 & 0 & \nu\gamma R_0 & 0 & 0 \\ 3\rho & -3\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3\rho & -3\rho & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3\rho & 0 & 0 & \epsilon\gamma R_0 & \gamma k_t R_0 & 0 \\ 0 & 0 & 0 & 3\rho & -3\rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3\rho & -3\rho & 0 & 0 & 0 \\ 0 & 0 & 3\rho & 0 & 0 & 0 & -\gamma & \lambda_r & \lambda_s \\ 0 & 0 & 0 & 0 & 0 & 3\rho & 0 & -\gamma-\lambda_r & -\lambda_s \\ 0 & 0 & 0 & 0 & 0 & 3\left(1-\xi\right)\rho & 0 & 0 & -\gamma-\lambda_r-\lambda_s \end{pmatrix} \tag{38}$$

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r^{\text{eff}}\chi_1 & \lambda_s^{\text{eff}}\chi_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r^{\text{eff}}\chi_2 & \lambda_s^{\text{eff}}\chi_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r^{\text{eff}}\chi_3 & \lambda_s^{\text{eff}}\chi_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_r^{\text{eff}}\chi_1 & -\lambda_s^{\text{eff}}\chi_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_r^{\text{eff}}\chi_2 & -\lambda_s^{\text{eff}}\chi_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_r^{\text{eff}}\chi_3 & -\lambda_s^{\text{eff}}\chi_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r^{\text{eff}}\left(\xi\chi_r + \left(1-\xi\right)\chi_{s,r}\right) & \lambda_s^{\text{eff}}\left(\xi\chi_r + \left(1-\xi\right)\chi_{s,r}\right) \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_r^{\text{eff}}\left(\xi\chi_r + \left(1-\xi\right)\chi_{s,r}\right) & -\lambda_s^{\text{eff}}\left(\xi\chi_r + \left(1-\xi\right)\chi_{s,r}\right) \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_r^{\text{eff}}\left(1-\xi\right)\chi_{s,r} & -\lambda_s^{\text{eff}}\left(1-\xi\right)\chi_{s,r} \end{pmatrix} \eta k_t R_0, \tag{39}$$

We obtained new estimates for k_t^{crit} in the cases of available TTI capacity and overwhelmed TTI capacity. Comparing these values to those obtained in our original model (with a single exposed compartment), we find that they do not deviate significantly. Graphically, we see that the distributions share the same properties, being slightly less skewed in the three-compartment case (see Fig. S6B). In numbers, the relative error we induce in calculating the critical level of contacts allowed for stability by using a single exposed compartment instead of three is reported in table S3.

Supplementary Table S3: Linearly-derived correspondence between contact reduction and the observed reproduction number (7.5e4 realizations).

		Percentile		
Variable		2.5 %	50 % (median)	97.5 %
$k_t^{\text{crit}}\big\vert_{\text{no TTI}}$	three compartments	37.80 %	42.58 %	46.68 %
	one compartment	37.71 %	42.59 %	46.71 %
$k_t^{\text{crit}}\big\vert_{\text{TTI}}$	three compartments	47.19 %	60.60 %	76.39 %
	one compartment	46.99 %	60.48 %	76.20 %
Relative error	no TTI	0.22 %	0.03 %	0.07 %
	TTI	0.42 %	0.19 %	0.24 %