A Workable Strategy for Covid-19 Testing:

Stratified Periodic Testing rather than Universal Random Testing[*]

Matthew Cleevely, Daniel Susskind, David Vines, Louis Vines, and Samuel Wills[**]

21 April 2020

Abstract

This paper argues for the regular testing of members of at-risk groups more likely to be exposed to SARS-CoV-2 as a strategy for reducing the spread of Covid-19 and enabling the resumption of economic activity. We call this 'stratified periodic testing'. It is 'stratified' as it is based on at-risk groups, and 'periodic' as everyone in the group is tested at regular intervals. We argue that this is a better use of scarce testing resources than 'universal random testing', as recently proposed by Paul Romer. We find that universal testing would require checking over 21 percent of the population every day to reduce the effective reproduction number of the epidemic, R', down to 0.75 (as opposed to 7 percent as argued by Romer). We obtain this rate of testing using a corrected method for calculating the impact of an infectious person on others, where testing and isolation takes place, and where there is self-isolation of symptomatic cases. We also find that any delay between testing and the result being known significantly increases the effective reproduction number and that one day's delay is equivalent to having a test that is 30 percent less accurate.

[**] Affiliations:
Matthew Cleevely, MEng MPhil MIET, founder of global online booking system 10to8 ltd.; matthew@cleevely.com.
Daniel Susskind, Fellow in Economics at Balliol College, University of Oxford daniel.susskind@balliol.ox.ac.uk.
Louis Vines; data scientist; louisgabrielvines@gmail.com.
David Vines; Emeritus Professor of Economics and Emeritus Fellow of Balliol College, and Senior Fellow in the Institute for New Economic Thinking (INET), University of Oxford, and CEPR; david.vines@economics.ox.ac.uk
Samuel Wills, External Research Associate, School of Economics, University of Sydney; OxCARRE, University of Oxford; CAMA, ANU; samuel.wills@sydney.edu.au.

1   Introduction and Summary

Governments around the world are looking for a testing strategy for Covid-19. And they are keen to see how this testing strategy might help ensure that the escape from lockdown is as speedy as a possible.

In a talk given on 3 April 2020, Paul Romer proposed population-wide testing and isolation – what we call "universal random testing".[1] He made the important points that the economic benefit of a speedier recovery would be measured in $trillions and this would easily justify spending $billions on testing, and that there is no necessity for such testing to be highly accurate. Unfortunately, as we will show, the model used in that talk is flawed. Our corrected model suggests that Romer's proposal is unlikely to work in practice. Instead, we show that targeted testing of particular groups, what we call "stratified periodic testing", and the subsequent isolation of those who test positive, would be a more effective tool in reducing the spread of Covid-19.

Romer uses a simple model to show how random testing of 7 percent of the population per day for evidence of infection (using antigen tests) would be sufficient to halt the pandemic. While already a big ask, we argue that his proposal ignores asymptomatic cases and rests on a mistaken calculation. A corrected model implies that the required proportion of daily testing would be more than 21 percent of the population – an impossible task – and indicates that attempting to test the entire population at random would be a waste of resources.

Instead we recommend testing smaller stratified specific groups who are at particular risk of spreading the infection. These would include healthcare staff, key workers, and others who are at high risk of creating cross-infection (See Section 6 below). Like Romer, we believe that high frequency testing of such groups would be manageable. But, unlike Romer, we also believe that only targeted testing would be feasible, and that, on its own, such testing would be a means of stopping the virus from spreading in the whole population.

We also recommend that testing of at-risk groups should be done periodically, rather than randomly. This ensures that each member is certain to be tested at regular intervals. This improves the efficiency of testing: for example, it can prevent some individuals being tested on two successive days, and others having to wait a disproportionate length of time to be re-tested. We show that this increases the efficiency of any given number of tests by approximately 35 percent, at no extra cost.

So long as testing remains a scarce and relatively expensive resource, we argue that testing of the general population should be reserved for immunity tests (antibody tests), which would allow those that have been infected to get back to work.

---

[1] See https://bcf.princeton.edu/event-directory/covid19_04/.

The paper is organised as follows. In Sections 2 and 3 we use our corrected model to calculate the frequency of random testing that would be required to halt the spread of the virus. We use our model to suggest that Romer's strategy would not halt the spread of the virus and that it is best thought of as a risky "throttle" strategy. In Section 4 we examine the robustness of these conclusions.

Section 5 sets out what we think is a workable testing strategy: testing people at regular intervals in carefully selected groups which have a higher rate of spread of the infection relative to others. This would enable greater rates of isolation in these groups, preventing the epidemic spreading where it matters most. We outline a number of practical suggestions that explain how these groups can be identified. And then in Section 6 we show how to formally calculate the required testing rate to sufficiently reduce the spread of the infection in these groups, when testing is done periodically rather than randomly.

Section 7 presents our conclusions.

In Appendix 1 we set out how and why we think that Romer made his error. Appendix 2 provides a simple formula which can be used to support the conclusions in the body of the paper that Romer's strategy would not control the spread of the virus, and which also provides some simple intuition as to what would be required to achieve such an outcome with universal random testing.

We should make it clear immediately that our paper is not intended as a criticism of Romer. His focus on the critical importance of testing, the ideas presented in his lecture of 3 April 2020, including his emphasis on the need to test key groups very frequently, and his simulations of the spread of the epidemic (which we discuss below) are all immensely valuable. We just want to "check the math" – as Romer has asked us all to do[2] -- and, by doing so, we demonstrate that doing the calculations properly shows that universal random testing should *not* be part of a Covid-19 testing strategy.

2   The basic ideas

The reproduction number R in an epidemic is the expected number of cases directly generated by any one infectious case. The *basic* reproduction number, R_0, is the initial value of R.[3] For Covid-19, Romer takes this value for the whole population to be R_0 = 2.5 and so will we.

---

[2] See Paul Romer's tweet, https://twitter.com/paulmromer/status/1248719827642068992.

[3] One can think of an *original* value for R_0 which is the reproduction number of the infection where all individuals are susceptible to infection, and no policy interventions have been adopted. But of course we also want to allow for the effects of policy interventions which might reduce that rate, such as wearing face masks, undertaking social distancing, or allowing the infection to spread so as to reduce the susceptible population. Such features can be included on our model. One way of doing this is to model these effects explicitly, which is

We know that an epidemic can only be controlled if the value of R is brought below 1. When that happens, each infected person infects less than one new person, and the epidemic will die out.

Romer wants to get the *effective* reproduction number R' (R prime) down to $R' = 0.75$ from the basic reproduction number of $R_0 = 2.5$, by randomly testing a fraction of the entire population each day and then isolating those who are found to be positive. In Romer's analysis, which we follow, the effective reproduction number R' is the product of the basic reproductive number, R_0, and the fraction of the infectious population that is not isolated. R' is below R_0 because a fraction of the population is tested each day and those found to be infectious are isolated.

Let φ be the proportion of the infectious population which is isolated. Then we can write Equation (1):

(1) $R' = (1 - \varphi)R_0$

For Romer's objective of $R' = 0.75$, Equation (1) implies that that $\varphi = 0.7$, *i.e.* that 70 percent of the infectious population is isolated, and that sufficient tests and isolation are carried out to make this possible.

Drawing on his calculations, Romer suggests that this can be done by randomly testing 7 percent of the population each day, *i.e* testing everybody randomly, at a frequency of about once every two weeks. He believes that this would achieve $R' = 0.75$. With a population of 300 million in the US testing on this scale would require about 20 million tests a day. In the UK with a population of 60 million, this would require about 4 million tests a day. Romer proposes the immediate allocation of $100bn in the US to make such an outcome possible.

This calculation, though, is mistaken. In Appendix 1 we identify the error in Romer's analysis of the relationship between the level of testing that is required to achieve a particular level of φ, and explain how we think that Romer came to make his mistake.

We now set out our alternative calculation which shows that, if random testing followed by isolation were adopted, more than 21 percent of the population would need to be tested each day. This, we show, is what would be necessary to get to a position in which 70 percent of the infectious population were isolated (*i.e.* $\varphi = 0.7$), so that $R' = 0.75$. To do this would, we show, require everyone in the population to be tested about every five days.

---

how we will allow for the effects of self-isolation. Alternatively one can be implicit, by using an *adjusted* value for R_0 as an input to the calculations, in order to reflect the characteristics of a population, or group of individuals, who have particular attributes – like a high proportion of susceptible people, or indeed recovered people - or are subject to already existing interventions - like lockdown.

3   Calculating the required test rate for universal random testing

3.1   Finding the testing rate which would get R down to *R'* = 0.75

We assume, like Romer, that the whole population is randomly tested. We let *t* be the proportion of the population tested each day, *i.e.* the probability, for each person, of being tested each day and then isolated. Romer allows for false negatives in his tests. He lets *n* be the proportion of false negatives and assumes that this is 0.3 (i.e. 30 percent of those who are infected test negative). We follow him in assuming such a high number.[4] In Section 4 we look at the effect of different proportions of false negatives.

We assume that the number of days an infected person is infectious is *d* = 14. Fourteen days is familiar in the analysis of Covid-19 as the period after which the person is either dead or – much more likely – recovered, but no longer infectious. Of course, this number may not be the best one to use. We use *d* = 14 here mainly because it is the number used by Romer as the number of days that each person who tests positive is placed in isolation. We discuss this issue further in Section 4 below and discuss Romer's procedure in Appendix 1.[5]

As a preliminary, let us consider the impact of an infectious person on others. We assume that if R_0 = 2.5, and if there were no testing which led to the isolation of infected people, then an infectious person would infect 2.5 people in total, or 2.5/*d* persons per day for *d* days. That is what we take R_0 = 2.5 to actually mean. Note that we discount any idea of a person being more or less infectious during the period of *d* days. See a brief discussion of this point in Section 4 below.

We now consider the impact of an infectious person on others when he or she has a probability *t* of being tested each day, and so of being placed in isolation immediately if the test is positive.[6] For clarity, it is helpful to think about the 'discovery rate' *x*, where *x* = *t*(1-*n*) and *n* is the number of tests which show false negatives, which we take to be *n* = 0.3 as specified by Romer. The variable *x* shows the probability that, on any day on which this person is infectious, he or she is isolated. This means that (1 − *x*) is the probability that this person will not be isolated, and so be able to infect people on the next day.

---

[4] There is a good reason for this. Massive testing – even of the amount which we contemplate – may well make it impossible to ensure that tests are accurate. Romer rightly argues that, whilst accurate tests are absolutely necessary for clinical reasons when treating an individual person, much more rough-and-ready testing is satisfactory if the purpose of this testing is epidemiological control.

[5] 7 days has been the standard advice for isolation or 14 days if more than one in a household, however recent data shows that the infectious period may last much longer. A recent detailed study of repeatedly tested individuals in Taiwan found a long tail for infectiousness. For further discussion of this point See Section 4 below. See also https://focustaiwan.tw/society/202003260015.

[6] This testing regime assumes that on the day you are tested, if tested positive and isolated, you cannot infect someone else. This simplifying assumption can be thought of in practical terms as a rapid 'early morning' test. In the section below on robustness we investigate the effect of a much more cautious assumption supposing that the infected individual remains infectious for the entire day that they are tested, with corresponding increase in required testing rate *t* for any given *R'*.

In almost all countries, an important part of the current policy response to Covid-19 is that those who show symptoms are required to self-isolate. Because of this, only asymptomatic people and those who have chosen not to self-isolate, for example with mild or mistaken symptoms, will be spreading the disease, and will be included in those who are tested. Romer discussed this issue in his lecture but did not actually allow for it in his calculations.

The proportion of asymptomatic patients isn't really known, and estimates vary wildly. The WHO suggests that as many as 80 percent of cases are asymptomatic or mild.[7] Very inexact data from Iceland suggest that all infectious patients are asymptomatic for the first five days and that, after this time, only about half become symptomatic.[8] We will use these Icelandic figures when constructing the "base case" in our calculations.

We construct our analysis as follows. We want to find the value of *x* that would give Romer's desired value for *R'* of 0.75. We will then work out the required probability of testing *t*, given that we take as given Romer's assumption that the proportion of false negative tests, *n*, is equal to 0.3.

If there were *no* self-isolation of those who became symptomatic, then the probability of an infectious person remaining undetected on day *j* is $(1-x)^j$ and the expected number of infections caused by an infected person on day j is r_j. Therefore, in expectation, an infected person would infect $r_1(1-x)$ persons on day 1 of their infection $r_2(1-x)^2$ persons on day 2, and so on, up to $r_{14}(1-x)^{14}$ persons on the final day *d*. We assume that the infection rate is constant over the infected period and so $r_j = R_0/d$.

We allow for self-isolation of those who are symptomatic as follows. We suppose that all of those infected are asymptomatic for five days (and are subject to random testing during that time). But that, after 5 days, a proportion of the population display symptoms and self-isolate.

Let α be the proportion of population who display symptoms and self-isolate. This means that only a proportion $(1 - \alpha)$ go on being tested from day 6 onwards. We can thus write our key equation as follows:

(3a) $R' = [(1-x) + (1-x)^2 + \ldots. (1-x)^5 + (1-\alpha)\{(1-x)^6 + (1-x)^7 \ldots + (1-x)^d\}]/d$

or, in more mathematical notation:

(3b) $R' = \frac{R_0}{d}\sum_{j=1}^{d}(1-x)^j - \frac{\alpha R_0}{d}\sum_{j=6}^{d}(1-x)^j$

---

[7] See https://www.medrxiv.org/content/10.1101/2020.02.20.20025866v2.
[8] See https://edition.cnn.com/2020/04/01/europe/iceland-testing-coronavirus-intl/index.html.

Our ambition, like Romer's, is to get to a position in which each infectious person on average only infects 0.75 other people, so that the virus will die out. So, we seek to find the value of x which would the total number of infections caused by such a person to 0.75. Recalling that $R_0 = 2.5$, and letting $d = 14$ we then can solve for x from Equation (4):

(4a) $R_0 [(1-x) + (1-x)^2 + \ldots + (1-x)^5 + (1-\alpha)\{(1-x)^6 + (1-x)^7 \ldots + (1-x)^{14}\}]/14 = 0.75$

or, in more mathematical notation,

(4b) $\frac{R_0}{d}\sum_{j=1}^{d}(1-x)^j - \frac{\alpha R_0}{d}\sum_{j=6}^{d}(1-x)^j = 0.75$

The solution to this equation can be obtained numerically for various values of α. When $\alpha = 0.5$, as roughly observed in Iceland, $x \approx 0.146$.

But $t = x/(1-n)$. And n, the proportion of false negatives, is equal to 0.3. So the required probability of testing on each day, t, is given by $t = 0.146/0.7 \approx 0.209$. That is to say, the proportion of people tested on any day must be as high as 21 percent in order to achieve the required 15 percent discovery rate x.

Recall from Equation (1) that $R' = (1-\varphi)R_0$, where φ is the proportion of the infectious population which is isolated. With $R_0 = 2.5$ and $R' = 0.75$, this means that $\varphi = 0.7$. Thus, over the 14 days in which a person is infectious, this person will, on average, be in isolation for 70 percent of the time. We have shown that to achieve such a very striking outcome, the probability of testing an infected person who is not yet isolated on any day must be at least 21 percent. With random, population-wide testing, this means, in effect, that everybody in the population has to be tested about every five days.

3.2   Identifying the testing "threshold" at which R' = 1

The dotted line in Figure 1 plots R' as a function of the proportion of the population tested. As we have already seen, when half of those who are infected self-isolate from day 6 onwards, *i.e.* $\alpha = 0.5$, t needs to be equal to about 21 percent to get R' down to 0.75. More than this, the Figure also displays the different testing rates which are required to obtain a range of different values for R'. And it does this for different values of α as well.

Figure 1
Over 20 percent of the population needs to be tested to stop the infection if no-one self isolates when they show symptoms
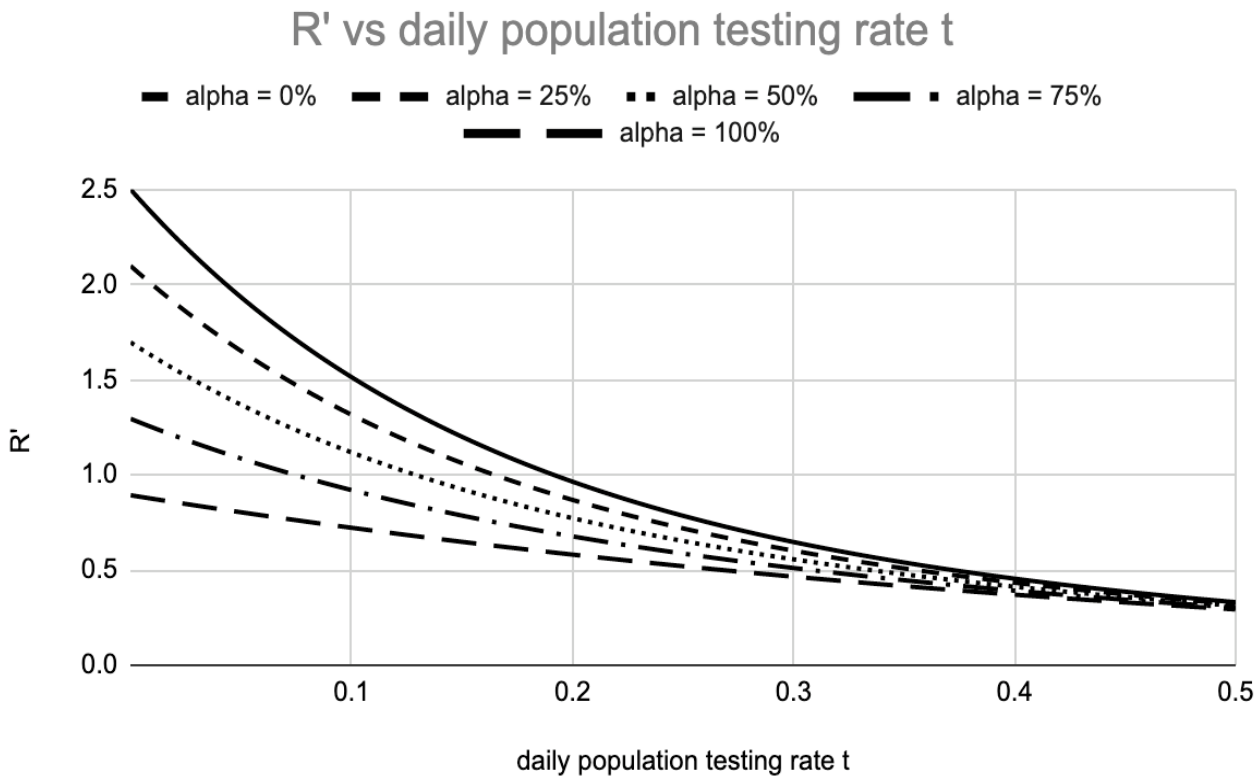
R' vs daily population testing rate t

- alpha = 0%   - - alpha = 25%   ·· alpha = 50%   - · alpha = 75%
- alpha = 100%

α is the proportion of infected people who display symptoms and self-isolate. Higher levels of α mean that the rate on infection (R') is lower for a given level of testing.

Figure 1 enables us to identify the threshold testing rates, t^* that reduce R' to exactly 1, and so just stop the epidemic from exploding.[9] For $α$ = 0.5, this threshold testing rate is t^* = 13 percent. To achieve this, everyone would need to be tested, on average, every eight days.

---

[9] It would be good to find a simple way of calculating the threshold testing rates, t^*, for any population, based on the value of R_0 for that population, and given any assumed value for $α$, without having to solve the complex non-linear model being discussed in this Section.

It turns out that we can do this by using a simple approximation that ignores the dynamics of the infection process, thereby producing an equation which is easy to solve. In Appendix 2, we set out this simple approximate method for calculating the threshold testing rates t^* for a population, depending on the values of R_0.and of $α$ that are appropriate for that population. We show that the approximation is relatively accurate, even although the calculation rests on two simplifying assumptions.

It might be useful to carry out the kind of calculations described in Appendix 2, even despite the fact that, in this paper, we are recommending using periodic testing rather than the random testing being discussed here. This is because, as we show in Section 5 of the paper, periodic testing is more efficient than random testing. As a result, calculating the amount of testing needed to bring R' down to 1, if testing were to be random, might well provide a *lower bound* for the threshold testing rate when using periodic testing. It might be useful to have a simple way of calculating this lower bound, even although the method of calculation is only an approximate one.

This is still way above Romer's proposed testing rate of 7 percent. Figure 1 also displays the sensitivity of the results to various values of α, the proportion of infected people who display symptoms and self-isolate. It shows that, at one extreme, when all cases are symptomatic after 5 days and then self-isolate (*i.e.* when $\alpha = 1.00$), the situation seems just about manageable: the epidemic would stop exploding even without any testing. Nevertheless, the required value of t which would ensure that $R' = 0.75$, is about 8.2 percent. This percentage is actually a little *above* that proposed by Romer because infected asymptomatic people do a lot of damage in the first five days! At the other extreme, with $\alpha = 0.00$, the situation is much worse: the probability of testing per day, t^*, which is required to stop the epidemic exploding is now nearly 20 percent (19.1 percent) and the value of t required to ensure that $R' = 0.75$ is about 26 percent (26.1 percent).

Data for the extent to which infectious people become symptomatic is extremely unreliable and furthermore the range of possibilities seems very wide. Even if over half of infected people self isolate, around 20 percent of the population would need to be tested every day to get R' to 0.75. The outcomes depicted in Figure 1 suggest that, on the balance of probabilities, Romer's strategy of universal random testing would be unworkable.

3.3   Romer's strategy is really a risky "throttle" strategy

Returning to the share of asymptomatic cases observed in Iceland, (*i.e.* the case when $\alpha = 0.5$), Figure 1 shows that by setting a testing rate of 7 percent of the population, the epidemic remains explosive with R' equal to about 1.3. In this situation, infection would spread rapidly in the earlier stages, since the testing rate is not high enough to slow the spread in a controlled manner; all testing would do is slow down the inevitable spread of the disease. However, ultimately, once the contagion reaches a certain size, the effect of testing, together with the fact that more and more people have had the disease and so are immune, will begin to slow the spread. The proportion of those infected will tend towards a constant level at which there is what has come to be called "herd immunity". It can be shown, for any R, that this proportion is given by $(1 - 1/R)$. Without testing, with a value of R_0 of 2.5, the herd immunity proportion is 60 percent.

Romer's strategy, with a massive amount of testing, would reduce R' to 1.3, and so, using the above formula, it would reduce the herd immunity proportion, to which the population is tending in the long run, to about 23 percent. But in the earlier stages of infection, when there is no immunity, the disease could still spread rapidly as the testing rate would not be high enough to slow the spread in a controlled manner. Thus, in sum, we can say that Romer's testing strategy would slow the spread, and would reduce the level of herd immunity, but would not control the initial explosive phase of the epidemic.

In fact, Romer is really proposing a random testing strategy which could be used as a throttle to control an inevitable spread, along with a view that, at 7 percent testing, this throttle strategy might be 'good enough'. But such a control mechanism does not stop very

large numbers of people being infected, it merely "flattens the curve". Of course, the hope is that something else intervenes, like a vaccine.

Romer has posted a very detailed and helpful model of the spread of the epidemic in this manner, one which avoids the problems of the model put forward in his April 3 talk. See https://paulromer.net/covid-sim-part1/ and https://paulromer.net/covid-sim-part2/.

In Romer's simulations of this model, the spread is very rapid: there is a peak of infections of between 3 percent and 18 percent of the population. This would overwhelm any national health system since even 3 percent of the population is an enormous number. There is also a chance that 20 percent of the population would be infected at the same time. That would be a national calamity.

Romer shows in his simulation model that using such a strategy over the course of 500 days might result in about 30 percent of the population contracting Covid-19. This is a risky strategy. Peak levels of infection might rise out of control. Even if testing rates were then increased, lags in responses would mean that the spread of infection would only be gradually reduced. Meanwhile, the virus would go on spreading towards the herd immunity level.

4   The robustness of our conclusion that universal random testing is unworkable

Of course, there are many changes to our assumptions which could modify our calculations.

In particular, there would be significant *reductions* in required testing rates if whole households were to be self-isolated if anyone in the household tested positive. If, for example, only one person in a household were tested at any time, and households consisted on average of two people, then each positive test would remove two people into isolation. That is – testing would become more effective.

On the other hand, our calculations have deliberately assumed a very speedy testing strategy: we have supposed that a test done on any day which finds the person to be infectious causes that person to immediately self-isolate even on that day; an extreme assumption. We have redone the calculations using the more cautious assumption that a positive test result for a test performed on any day does not lead to the person isolating until the next day. The relevant equation now becomes:

(5a) $R' = R_0[1 + (1 − x) + (1 - x)^2 + …. + (1 − x)^4 + (1 - \alpha)\{(1 − x)^5 + (1 − x)^6 …. + (1 − x)^{13}\}]/14$

or, in more mathematical notation.

(5b) $R' = \frac{R_0}{d}\sum_{j=1}^{d}(1 − x)^{j−1} − \frac{\alpha R_0}{d}\sum_{j=6}^{d}(1 − x)^{j−1}$

These results are *much* worse than those described in Section 3. With α = 0.5, the critical testing rate, t^*, is now 16 percent, and the testing rate required to get R' down to 0.75 is now as high as 27 percent.

Furthermore, we have been assuming, like Romer, that there is uniform contagion throughout the 14-day period. But the medical data shows an asymptomatic infectious period followed by a hump of maximum infectivity as symptoms develop and a tail as symptoms resolve.

We conclude that for the random universal testing proposed by Romer to be workable (involving testing, say, less than 10 percent of the population per day) policymakers would require confidence that:

i) nearly all infected patients are symptomatic and self-isolate, reducing the burden on testing after the incubation period,

ii) the tests are sufficiently effective, and complied with, that they capture more than 70 percent of infected cases (and ideally close to 100 percent),

iii) testing is conducted quickly, early in the morning, and people are isolated on the day of the test, and

iv) whole households are isolated when any member is infected.

Unfortunately, we do not feel that all these conditions can be met given our current state of knowledge about the virus so we do not believe that whole-population random testing would be a good use of resources.

5   A workable strategy of stratified periodic testing[10]

5.1   Stratified testing

We argue that testing should be carried out at different frequencies for different stratified groups, based on their likelihood of infecting others. This likelihood can be deduced from their occupation, geography, and other factors. Testing at rates above 20 percent per day could be done for carefully selected groups which have a high basic reproduction number (R_0) relative to others. This would enable greater rates of isolation in these groups, lowering their effective reproduction number (R') and helping to prevent the epidemic spreading where it matters most.[11] This appears to be a much lower-risk strategy to contain the spread of infection, and could be done with cheap tests, even if they are somewhat inaccurate.

---

[10] Some of what follows comes from suggestions made to us by Eric Beinhocker, for which we are very grateful.
[11] It might even enable the general lockdown to be eased, so that other lower-risk groups could keep working and not need to be isolated.

Broadly, there are two types of people that are likely to have a particularly high basic reproduction number relative to others. The first are those who have a high basic reproduction number to begin with. These are individuals who would have been more likely to infect others, before the infection had begun to spread and any policy interventions had been adopted. Doctors are a one example. They have very frequent and unavoidable close contact with others. Their basic reproduction number will, as a result, be very high and, very frequent testing will be necessary to ensure that their effective reproduction number is low enough. As a result, there could be very frequent testing for doctors in hospitals to ensure that the effective reproduction rate for them is brought well below 1. It appears likely that the relevant calculations will show that doctors actually need to be tested every day. This is something which Romer already suggests in his talk.

But there are also other people that will have a high basic reproduction number because of the uneven application of the lockdown and other factors that might affect variation in infectiousness across groups.[12] For many people, for instance, lockdown means that they are confined to their homes (including many workers who are able to work from home), reducing their basic reproduction number well below 1. But key workers, who are encouraged to keep working in spite of the lockdown, will have a higher basic reproduction number as a result (*e.g.* those involved in food production and distribution). The same will be true for all those who are unable to work from home and are given permission to avoid the lockdown (*e.g.* those involved in construction and manufacturing). Another group that is likely to have a higher basic reproduction number involves those who are more exposed to people who are particularly susceptible to the infection (*e.g.* prison warders and care workers). As testing is rolled out it will become increasingly appropriate to frequently test all such people.

One challenge in all this is that the basic reproduction number may be high in particular groups for idiosyncratic reasons that are hard to anticipate. The kinds of calculation described in the next section could easily be carried out for structured samples in different locations, in order to identify these pockets of infectiousness.

The general principle, then, is that testing should be concentrated in groups that have high basic reproduction numbers relative to others. But this principle should not be interpreted too strictly. In certain cases, other criteria may also be important: for instance, we may want to regularly test groups that interact with those who are more likely to die from the infection (this is another reason to test care workers more frequently) or groups whose absence would have a greater economic impact than others (this is another reason to test key workers more frequently).

---

[12] See footnote 3.

5.2   Clarifying how our strategy differs from that of Romer

We hope that Romer would agree with what we have just written above. Indeed, in the version of his plan that he set out on Twitter, Romer has himself provided useful suggestions about who might have priority as tests are rolled out.[13] But from here on we part company.

Romer goes on to suggest that, once tests have been rolled out for these most important groups, there be a further vast expansion of testing, enabling mass random testing *for the whole population,* in order to get the effective reproduction number down *for the whole population.* The version of his plan on Twitter makes this very clear. It concludes as follows:

> "When you strip away all the noise and nonsense, note that once we cover essential workers, it's easy to test everyone in the US once every two weeks. Just do it. Isolate anyone who tests positive. Check your math. Surprise, R_0 < 1. Pandemic is on glide path to 0. No new outbreaks. No need for any more shutdowns."[14]

Instead, we argue that testing must be focused on particular groups. This is because our findings, discussed in Section 3 above, show that a mass testing plan would still leave the effective reproduction number significantly above 1 unless it was carried out infeasibly frequently.

Nevertheless, there will still need to be random testing of groups in the population, and some random testing of the whole population. But this testing would be for *informational* purposes only and would only involve testing very small samples of those involved.

Such informational testing will be needed for two reasons. First, random testing of small samples from particular groups will be necessary be to track groups in which the basic reproduction number is already known to be high, and where there is greater potential for a high rate of spread. Once identified, these groups will then need very frequent testing of everyone in the group, for the reasons which we have been discussing in this paper. But testing of small samples of wider groups in the whole population will also be needed to identify new groups where the basic reproduction number is high. As before, this may be for idiosyncratic reasons that are hard to anticipate. Once identified, such groups will then also need very frequent testing of everyone in the group.

But the accuracy of this testing for informational reasons will be determined by the sample size, rather than population size. The samples required for these informational purposes will be *very* small relative to the size of the whole population.

---

[13] See: https://threadreaderapp.com/thread/1248712889705410560.html.
[14] Romer uses R_0 here to stand for what we call R'. See again:
https://threadreaderapp.com/thread/1248712889705410560.html

5.3 Periodic tests rather than random tests

Once the frequency of testing has been decided and testing kits are available, testing can begin for everyone in the identified groups. But it is important that this testing be done periodically for each person, rather there being a random choice of those who are to be tested in each time period.

The rationale underlying periodic testing can be explained by the "waiting-time paradox".[15] Random testing, say of 20 percent of a group each day, wastes many resources. This is because every day some of those tested will have actually been tested the day before, whilst others of those who are infectious will, nevertheless not be tested and so will possibly continue infecting people. By contrast, periodic testing of 20 percent of a group means that, on days 1 to 5, a different fifth of the group will be tested each day, and that on day 6 the first fifth of the group will tested again, and so on. It is clear that this means each person tested will have been tested exactly five days previously, removing the problem that some tests are being wasted and that other tests are being postponed for too long. Because of this, you need to test far fewer individuals in a group to get the same reduction in *R*.

In the next Section, we provide a simple account of this issue, and show how important it is likely to be. We show that with high testing rates, periodic testing beats random testing by a very significant factor. For example, in the model which we examined in Section 3, in the special case in which there is no self-isolation of symptomatic people (*i.e.* with $\alpha = 0$) and perfect testing (*n* = 0), our testing rate required for *R'* to be 0.75 was 18.3 percent with random testing. With periodic testing this rate falls to 13.5 percent, a 26 percent reduction. This is a big improvement at *no* extra cost.

5.4 The testing system: running two kinds of tests in parallel

In this paper, we have been discussing antigen testing (i.e. testing for active infections) as opposed to antibody testing (i.e. testing for those who have had the disease and are both immune and non-infectious). A combination of the two might be effective and realistic if the testing capacity for active infection remains constrained, but that one-time antibody tests become widely available. One might then proceed as follows:

- Immediately and frequently perform antigen tests on groups and areas with a high R' and immediately isolate those found to be positive. Trace[16] and test the contacts of

---

[15] The waiting time for a Poisson bus service is twice the waiting time for a periodic bus service with the same rate for a randomly arriving traveller.

[16] Contact tracing can be as simple as testing those in the household and workplace of those who are infected. Technological solutions exist to perform more detailed contact tracing, *e.g.* using mobile phone movements. However, these involve privacy concerns which, crucially, may take time to debate and resolve. The authors believe that simple, immediate testing of infected households and workplaces is preferable to detailed tracking of mobile phone movements at some months delay. As contact tracing apps become more

those who test positive, as these now have a higher probability of also testing positive.

- Self-isolate anyone developing symptoms for a minimum of 7 days. These people would not be tested unless medically necessary. Trace and test their contacts.
- If there were enough tests then one could test people at the *end* of their isolation period to show that they were clear of virus before they were allowed to come out of isolation.
- Widespread home kit antibody testing for anyone to see if they had had the virus - these would be one-off tests that would not need to be repeated.
- A system to track people with immunity who could then circulate freely if they had either a) had a positive antibody test, b) had a positive active infection test more than some specified number of days ago, or c) had a negative active infection test after their symptoms resolved.

All of this could be done using cheap antigen tests, even if they were somewhat inaccurate. It is not the case that "no test is better than an unreliable test". Our calculations show that, whilst accurate antigen tests are absolutely necessary for clinical reasons when treating an individual person, much more rough-and-ready testing is satisfactory if the purpose of this testing is epidemiological control through isolation. (In our baseline calculations discussed below, for instance, we assume that 30 percent of infected people wrongly test negative, $n = 0.3$).

Doing all of this will help governments to track spread and to determine where hotspots are flaring up. Such information will help them to work out how to selectively tighten, or loosen, containment measures when needed.

Such a testing procedure would involve doing two things at once: the stratified periodic antigen testing which we have been discussing would be designed to damp the spread of the disease in key groups, by catching those in these groups who were infectious but asymptomatic, or pre-symptomatic, or post-symptomatic, and so not self-isolating. At the same time, antibody testing for the entire population would separate out the immune population; passing an antibody test would enable such people to return to work.

6   Calculating the required test rate with random testing[17]

6.1   Finding testing rate which would get R down to R' = 0.75 using periodic testing

In this Section we show how to solve for the required testing rate when there is periodic testing. As in our discussions of random testing in Section 3 we aim to find the testing rate which would get R down to R' = 0.75, and also to Identify the testing "threshold" at which R'

---

widespread the degree of contact (separation distance, length of interaction, *etc.*) will be available to help target testing resources.

[17]   We are grateful to Frank Kelly for his assistance in preparing this Section.

= 1. One of our aims is to show how much more effective periodic testing might make the testing process, when compared with the random testing process discussed in Section 3.

As in that discussion of random testing, our first objective was to get the value of R down from R_0 = 2.5 to R' = 0.75. In our discussion here, we will include the effects of self-isolation, *i.e.* the results in Section 3 with which we will compare our findings here are those in which α = 0.5. Our results suggest that periodic testing might be about 37 percent more effective than random testing, at no extra cost.

As in Section 3, we let R' be the expected number of people that a randomly chosen infected person infects before that person is positively tested (or stops being infective, if sooner). Let r_j be the expected number of individuals infected by an individual on day of his/her infection, for j =1,2,...,d where the length of infectivity is d. Thus $R' = \sum_{j=1}^{d} r_j$ Now suppose an individual is tested every N days, and that for high risk groups we test very frequently, so that $N < d$. If the time of the infection is random and the individual is infective for the day of the test (as we considered in more likely in our robustness analysis in section 4), then

(6) $\qquad\qquad R' = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{i} r_j$

From this we can deduce that

(7) $\qquad\qquad R' = \frac{1}{N}\sum_{j=1}^{N}(N - j + 1)r_j$

Here we introduce our self-isolation factor α into Equation (7) by noting that r_j during the infectious and asymptomatic period (first 5 days), which we denote as $d_0 = 5$, is still $r_0 = R_0/d$ and thereafter $r_\alpha = (1 - \alpha)R_0/d$ for the infectious period in which a person may self-isolate. Substituting these values into r_j we have three alternative situations to consider: (i) the case when $N < d_0-$ the testing rate is more frequent than the move to becoming symptomatic that occurs after day d_0 and therefore Equation (7) represents this outcome entirely, (ii) the case when $d_0 < N < d$ for which the testing period N is between 6 and 13 days inclusive, and finally (iii) the case when N ≥ d.

The following is for the case when d_0 < N < d which is the testing range that tends to deliver an R' < 1 for this model of testing

(8) $\qquad R' = \frac{1}{N}\sum_{j=1}^{N}(N - j + 1)r_j = \frac{1}{N}\sum_{j=1}^{d_0}(N - j + 1)r_0 + \frac{1}{N}\sum_{j=d_0+1}^{N}(N - j + 1)r_\alpha$

We take $r_\alpha = (1-\alpha)R_0/d$ to the daily likelihood an infected individual will infect another person during the period in which they may become symptomatic and self-isolate with probability α. This gives[18]

(9) $R' = \frac{R_0}{2Nd}[(N+1)N - \alpha(N - d_0 + 1)(N - d_0)]$

We can now compare our findings with those in Section 3, for the case in which α = 50 percent. We found there that that to bring R down from $R_0 = 2.5$ to $R' = 0.75$ would – with random testing and next day test results - require a testing rate of 27 percent of the population a day, (or about one test every four days) . But those results assumed that 30 percent of tests failed. If tests were perfect those results imply an equivalent correct identification rate of 0.19 per day (or a 100 percent accurate test about every five days).

We now use Equation (9) to solve for the value of N, the number of days between each test, that is required to bring R down from $R_0 = 2.5$ to $R' = 0.75$ when there is periodic testing. This equation shows that the period between testing for each individual would need a test rate of 17 percent or a test every 6 days[19]. This assumes a test that has the same 30 percent false negative rate to estimate the periodicity of the test. If these tests were perfect, then the test rate could be every 8.2[20] days or an identification rate of 0.12 - that is a test with a correct identification rate carried out on 12 percent of the population per day.

This is a thirty-seven percent reduction in the rate of testing required, as compared with the case of random testing. We can see that doing random testing would provide a big improvement at *no* extra cost.

6.2   Identifying the testing "threshold" at which R' = 1

The dotted line in Figure 2 plots R' as a function of the proportion of the population tested.[21] As we have already seen, when half of those who are infected self-isolate from day 6 onwards, i.e. α = 0.5, t needs to be equal to about 17 percent to get R' down to 0.75.

---

[18] For the case N < d_0 the solution remains unchanged as the frequency of testing is above that which would allow the self-isolation process to occur: $R' = R_0 \frac{N+1}{2d}$
and for the case $N \geq d$:
$$R' = R_0\left[1 - \frac{1}{2N}(d-1)\right] - \frac{\alpha R_0}{2dN}(d - d_0)(2N - d0 - d + 1)$$

[19] We make here a simplifying assumption that in order to reproduce an effective testing rate x given the false negative rate of a test n, that the required *t* is *t = x/(1-n)*.
[20] Of course, in reality such a number would need to be rounded up or down to a full number of days.
[21] It is interesting to note that when N<d_0 we are testing such that the virus spreads only for N days precisely (the time between tests) which means that R' changes linearly when we reach this high frequency level of testing.

More than this, the Figure 2 also displays the different testing rates which are required to obtain a range of different values for R'. And it does this for different values of α as well. Figure 2 enables us to identify the threshold testing rates, t* that reduce R' to at which the disease exactly 1, that is to the value which divides outcomes in which the epidemic dies out from outcomes in which it explodes

For α = 0.5, this threshold testing rate implied to bring R' to 1 is t* = 11 percent – testing 11 percent of the population each day. To achieve this, everyone would need to be tested every nine days. So with periodic testing of an entire population with good isolation of infected individuals, the population of infected individuals may maintain a stable size (R'=1).
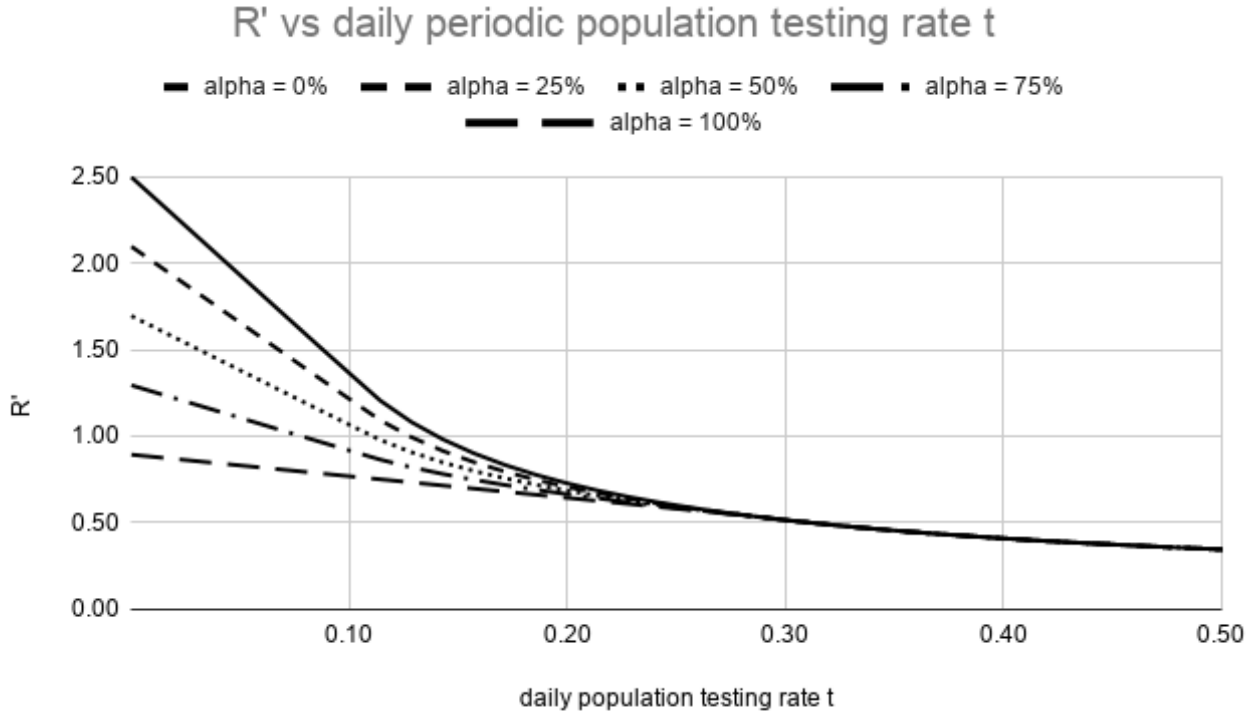
Figure 2 also shows us how sensitive this testing rate is to the effectiveness of a population's self-isolation when infectious. For testing rates that are below 20 percent per day (*i.e.* less frequent than once every 5 days), the period in which a person is assumed to be asymptomatic) we can clearly see the risk of those who are infectious not successfully self-isolating.

This is an important figure as *α* may change from population group to population group. For example, contract workers who are paid by the hour have a direct incentive to ignore symptoms. This would translate to a lower *α* for this group and consequently a much higher R' for any rate of testing. Workers in this category who also have high contact rates as part of their job would therefore be expected to require the highest rate of testing.

In the case in which all individuals become symptomatic and so self-isolate after 5 days (when *α* = 1.00) R' ≈ 1. This means that the pandemic doesn't necessarily die out and therefore to bring this down to zero, by bringing R' to 0.75 we need a 'perfect' (no false negatives) test every 12 days. As we show in the following section, if we have same day testing and a perfect population and test, we can reduce the testing rate to every 19 days.

Figure 2

With periodic testing the proportion of the population checked each day can vary depending on how likely it is that infected people self-isolate.

R' vs daily periodic population testing rate t

— alpha = 0%   – – alpha = 25%   ▪▪ alpha = 50%   — ▪ alpha = 75%
— — alpha = 100%

α is the proportion of infected people who display symptoms and self-isolate. Higher levels of α mean that the rate on infection (R') is lower for a given level of testing.

6.3   The impact of an instant test

In Section 3 we started by looking at a theoretically perfect test for which a positive identification of an infected individual would result in them not being able to infect anyone that day. Whilst this instantaneous test is unrealistic, it is useful to compare the results of our periodic model against those of the model in section 3.

We can solve the model introducing a delay with similar results[22]; Again, the impact of differing self-isolation likelihoods is prevalent when the testing rate falls much below 20

---

[22] To model the instant test we adjust our model to remote the anticipated extra day of infection that would have otherwise occurred, deprecating the sum used in earlier sections from j=1 to i, to j=− 1 to i-1. For completeness we now explicitly set out the equations for these calculations explicitly.

(a) For low-risk individuals we can test much less frequently, so that $N \geq d$. Then we obtain:

$$R' = \frac{1}{N}\left(\sum_{i=1}^{M}\sum_{j=1}^{i-1} r_j + (N-d)R_0 \right)$$

and hence

$$R' = R_0 - \frac{1}{N}\sum_{j=1}^{d} j r_j$$

So, if $r_j = R_0/d$ we can solve for the required rate of testing. It is

$$R' = R_0\left[1 - \frac{1}{2N}(d+1)\right] - \frac{\alpha R_0}{2dN}(d - d_0)(2N - d0 - d - 1)$$

(b) Likewise for testing when $d_0 < N < d$ we now sum to i-1 as opposed to i:

percent. However, as the tests are now instant, the amount of testing in order to reduce R has fallen. For α = 0.5 the required testing rate to achieve a value of R = 0.75 is now every 7.7 days (t = 13.4 percent) as opposed to every 6 days for a test that gave the results a day later. It is interesting to note that the effect of an unreliable test (30 percent false negatives versus no false negatives) is similar in scale to the impact of having to wait for a day for the results of a test (so that there is an additional day on which the person can spread the infection). The figures show an effect of 13 percent for an unreliable test versus 12 percent for a perfect test which would deliver results a day later.

Figure 3
If the results of the test are known instantly then less of the population needs to be checked every day to reduce R' to a given level

R' vs daily periodic population testing rate t with an instant test
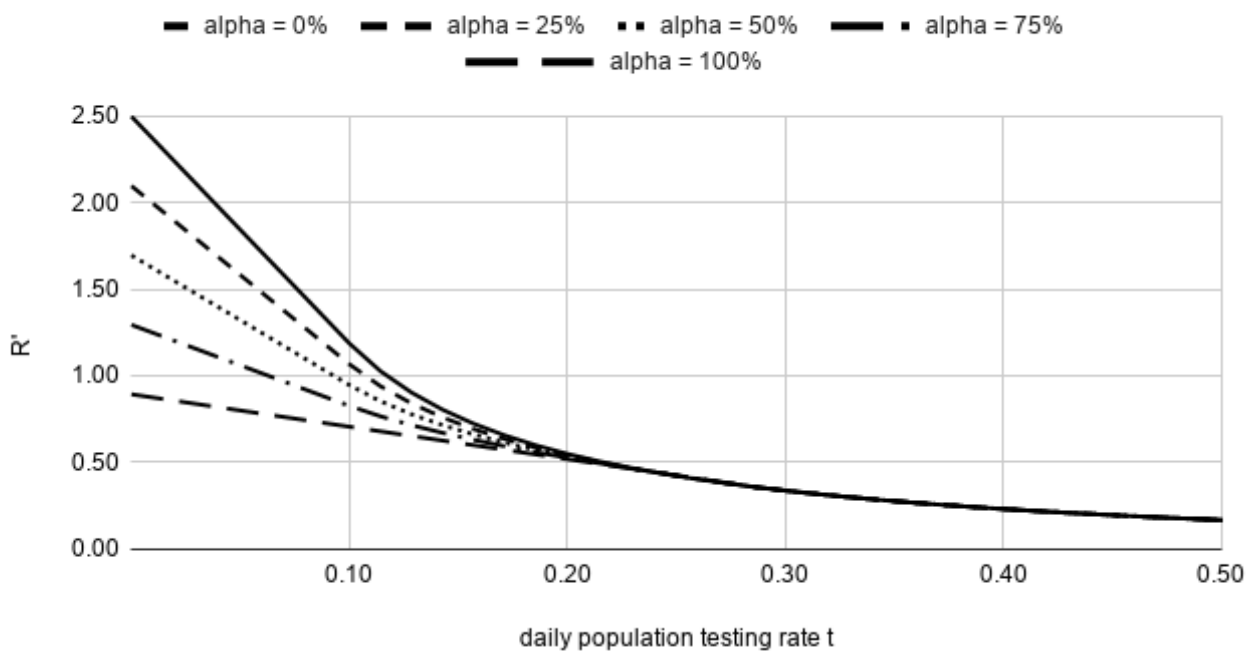
α is the proportion of infected people who display symptoms and self-isolate. Higher levels of α means that the rate on infection (R') is lower for a given level of testing.

$$R' = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{i-1} r_j$$

Then;

$$R' = \frac{1}{N}\sum_{j=1}^{N}(N-j)r_j$$

Giving:

$$R' = \frac{R_0}{2Nd}[(N-1)N - \alpha(N - d_0 - 1)(N - d_0)]$$

(c) For high frequency testing for when $d_0 > N$ and therefore α is not relevant to the dynamics:

$$R' = \frac{1}{N}\sum_{j=1}^{N}(N-j)r_j$$

$$R' = \frac{R_0}{2d}(N-1)$$

6.4   Allocating testing over populations with different R_0

If scarce testing is to be allocated over individuals with different prior probabilities of infection, these formulas can be used to optimize the allocation.

If we take the simplest form of our periodic testing model, when α = 0, for high frequency testing where an individual has a test after a smaller number of days than the length of an infection (N<d), then equation (7) simplifies to:

(10) $R' = \frac{R_0}{2d}(N+1)$

Suppose that individual k has prior probability p_k of infection and that we wish to choose N_k so as to allocate a given amount of testing over a set of individuals so as to maximally reduce R'. Then, using frequent testing where an individual has a test after a smaller number of days than the length of an infection (N<d), the optimal allocation of a given amount of testing should choose $N_k \propto 1/\sqrt{p_k}$.

This shows very clearly that there are diminishing returns to making N very small – that is conducting more frequent tests results in increasing but diminishing returns to R'. The amount of testing allocated, therefore, to higher risk individuals is naturally higher, but not proportionately so.

6.5   Summary

Periodic testing periods of around 6 days may be sufficient to control the propagation of infection on their own. This is testing approximately 17 percent of the population each day. This figure is highly sensitive to changes in self-isolation behaviour and the characteristics of the test.

This model moves us towards a useful framework for determining the frequency with which a population need be tested in order to treat that group's R_0. For any particular population, the frequency of testing which is required to reduce transmission sufficiently is determined by the both the population's properties and the properties of the test. A population's key factors are the capability of its individuals to self-isolate if symptomatic, and the initial propagation rate R_0. The important figures for the test being administered is the speed of results and the accuracy with which it reports positive results. All of these can be traded off against one another using our model.

7   Conclusions

This paper proposes 'stratified periodic testing' as a strategy for ending the lockdown and returning economies to work, while preventing an explosive re-emergence of Covid-19 (keeping R'< 1). The testing would be 'stratified', in the sense that it would focus on specific

subsets of the population who currently have the highest basic reproduction number R_0. The criteria could be amended to also take into account the vulnerability of subgroups, or the loss of economic activity if they are forced to self-isolate at home. The testing would also be 'periodic', in the sense that each member of the subset would be tested at regular, defined intervals, rather than testing within the group being done at random. This ensures that infected people can be identified and isolated quickly. Those who test positive would be quickly isolated at home, as would anyone with symptoms. The tests need not be perfect: if they are cheap but deployed widely within particular groups, then false negatives can be offset by the scope of testing. The effectiveness of the program can be improved by simple tracing of the contacts of those infected: for example testing those in their workplaces and households, rather than fully tracking mobile-phone movements with the associated privacy concerns.

We argue that this is better than 'universal random testing' which is currently being discussed globally. Romer suggests that by testing 7 percent of the population every day we can get the effective reproduction number of Covid-19 to around 0.75 and curb the epidemic. Unfortunately, these calculations contain errors. By correcting this method, and using reasonable assumptions about asymptomatic carriers, we believe that at least 21 percent of the population would need to be tested each day to get the effective reproduction number well below 1 (*i.e.* to the value of 0.75). For obvious reasons, we do not see this as a feasible population-wide strategy.

Any testing strategy should be thought of as a complement to other measures that can reduce the spread of Covid-19 at little economic cost. For example, those that can work from home with little loss of productivity should continue to do so, retirees should continue to self-isolate, and people in public places should wear masks and regularly wash their hands. Stratified periodic testing can then help those sectors in the economy that cannot operate from home get back to work quickly and safely. This should continue until widespread vaccines or treatment for the virus are available

Appendix 1: **Romer's analysis of testing**

We now explain why we think there is a mistake in the way in which Romer calculates φ, the proportion of the infectious population which is isolated. We then present our attempt to understand how and why he made his error.

A.1.1 Romer's analysis

Romer assumes random testing of the whole population. In his calculations, he lets t be the proportion of the population tested each day, *i.e.* the probability that each person is tested each day. He supposed that $t = 0.07$.

Romer allows for false negatives in tests. He lets n be the proportion of false negatives. Romer assumes that this proportion is 0.3.

Romer lets *l* be the number of days that each person who tests positive is placed in isolation. He assumes *l* = 14.

Romer then computes φ, the proportion of the infectious population which is isolated, as follows. He writes something similar to, but not the same as what we have called Equation 1 in our paper.

(1) $\varphi = t(1 - n)l$

Just to be clear, this equation here comes directly from the slides which accompanied Romer's talk.[23] We have no background on why he wrote down this equation, and we think that it is incorrect. We say this because in our paper above Equation (1) reads $\varphi = t(1 - n)$d. This has the variable d on the right-hand side, showing the number of days for which a person remains infections. By contrast Equation (1) above has *l* on the right-hand side the variable *l* which is the number of days that each person who tests positive is placed in isolation. The nature of what we think is Romer's error is discussed immediately below.

Since *t* = 0.07, (1 − *n*) = 0.7 and *l* = 14 Romer claims that φ = 0.69. This value of φ = 0.69 would, he says, produce his desired value for *R′*, since:

$R' = (1 - \varphi) R_0$, or $R' = (1 - 0.69) \times 2.5 \approx (1 - 0.7) \times 2.5 = 0.75$.

Drawing on these calculations, Romer suggests that there should be testing of 7 percent of the population each day.

Notice that, although Romer mentioned self-isolation of those who have symptoms in his lecture, there is no allowance for such an action in any of the calculations in his slides.

A.1.2 Our Criticism

It is helpful to try to understand how Romer made what we think is an error.

To see most clearly, and simply, why his calculation cannot be right, imagine what would happen if there were to be double the amount of testing proposed by Romer, *i.e.* suppose that *t* = 0.14. Then, using his formula for φ we would get $\varphi = t(1 - n)l$ = 0.14 times 0.7 times 14 = 1.38; the person would be in isolation for more than all of the period of 14 days! So the equation must be wrong.

How can we understand the inclusion of '*l*', the number of days that an infected person is placed in isolation, on the right-hand side of this equation? One possibility is that the inclusion of '*l*' is simply a mistake. Romer states that φ=t(1-*n*)*l*, but φ and t(1-*n*)*l* appear to be very different things. Because t(1-*n*) is equal to the probability that an infectious person is put into isolation on any day, it follows that t(1-*n*)*l* is equal to the expected length of isolation any infected person is likely to face, after one round of testing. But φ is the fraction of the infected population that are isolated -- which is clearly not the same thing as the

[23] See minutes 16 to 20 of the Romer talk, and the accompanying slides.

expected length of isolation any infected person is likely to face, $t(1-n)l$. So it appears that stating $\varphi=t(1-n)l$ is a mistake.

Another possibility is to make a set of assumptions about Romer's set-up that bring the meaning of φ and $t(1-n)l$ closer together. For instance, consider the following approach. First, interpret 'l' as the 'number of days that an infected person is infectious', rather than 'is placed in isolation'. Secondly, define 'Z' as the number of infected people not in isolation. Thirdly, imagine there are 'l' periods where, in each period, a fraction $t(1-n)$ of those infected people not in isolation, Z, are removed and put into isolation. And finally, assume that in each period the number of infected people not in isolation, Z, remains the same (i.e. the infected who are put into isolation are replaced with newly infected people). Then it follows that, after 'l', periods, $t(1-n)l *Z$ people will be in isolation. Now, $t(1-n)l$ is indeed equal to φ, the proportion of the infected population not in isolation who are put into isolation – but with two very significant caveats. First, it assumes that everyone who will be isolated over the 'l' days is isolated on the first day. And secondly, because Z is constant over time, it follows that φ may also be greater than one if t or l is large enough, or n is small enough – which, as shown before, is exactly the problem with Romer's analysis.[24]

Appendix 2   A simple method for calculating the testing threshold when testing is random

In this Appendix we set out a simple method for calculating the threshold testing rate, for random testing, which would reduce the effective reproduction number, R', to the value at which the disease does not die out. The calculation employs a simple approximation which ignores the dynamics of the infection process. If we ignore the dynamics, we do not have to solve a complex equation like Equation (3) which sums a number of effects in a non-linear way, over many time periods.

Our method of calculation builds on the following insight: for R' to be less than 1 when there is universal random testing then, on any given day, a person with Covid-19 is more likely to go into isolation than to spread it to someone else. Relying on this insight, we can ignore the dynamics of the process and simply solve for the value of t for which this condition will hold. We proceed in two steps.

(a)   For simplicity, we first examine the extreme case in which none of those who are infected become symptomatic and self-isolate; this corresponds to the case considered in Section 3 in which $\alpha = 0$.

---

[24] Intuitively, the problem here is that you are taking a fraction $t(1-n)$ of the infected population not in isolation, Z, and putting them into isolation in each period – but because Z replenishes over time, if you isolate a large enough proportion of Z, $t(1-n)$, enough times, l, then you will end up with more infected people in isolation, $t(1-n)l*Z$, than there are infected people not in isolation, Z.

Consider any group of z, as yet unidentified, infectious people. Assuming that this group is a small fraction of the overall population, the number of people who will be infected by this group on any given day is (R_0/d) times z, where d is the number of days that an infectious person remains infectious.

The number of these z people who, on this same day, will go into isolation because they have tested positive will be $t(1-n)$ times z. But there will be additional infectious people who cease to infect others because, although they did not test positive on that day, the period during which they had the disease and were infectious will have come to an end. This happens with probability $1/d$; so there will be $\{[1-t(1-n)]/d\}$ times z such people[25]. Thus, for R' to be less than 1, we require that:

(1) $\{t(1-n) + [1-t(1-n)]/d\} > R_0/d$.

This means that, for this extreme case, the threshold testing rate is given by

(2) $t^* = (R_0 - 1)/[(d-1)(1-n)]$

If $R_0 = 2.5$, $d = 14$, and $n = 0.3$, we get $t^* = 16.5$ percent . That is, this method says that, to get R less than 1 by randomly testing the whole population, one needs to test at least 17 percent of the population. That is, the threshold testing rate, t^*, is 17 percent.

This is a lower value than what we found for t^* using the full dynamic model in Section 3 when $\alpha = 0$. The result there was that $t^* = 19.1$ percent. The discrepancy between these two results arises precisely because of the dynamic process of the epidemic: the simple calculation carried out here ignores the fact that, as time passes, testing will remove some of the infected people, so that they are no longer available to be tested on later days. That is what made Equation (3) so complex.[26] For this reason the result produced using this method will always underestimate the required testing rate. This simple method thus provides a (quick and dirty) lower bound for the true value of t^*. Nevertheless, the fact that this calculation is so simple, and the intuition provided by thinking about the problem in this way, may make it useful to carry out this calculation.

(b)   This calculation can be readily extended to include the more general cases considered in Section 3 in which a proportion of those who are infected become

---

[25] This $1/d$ probability is the chance that an infected person becomes non-infectious independently of testing. An intuitive way to think about this is that, in choosing someone at random, there is a $1/d$ chance that that person is on their last day of infection and so will become non-infectious the following day. This is only an approximation since it requires that the value of R' resulting from testing is equal to 1. That is because if $R'>1$ then the virus would be spreading and hence an individual would be less likely to be on their last day of infection; conversely if $R'<1$ then a higher proportion of the infected population would more likely to be about to end their infectious period.

[26] It is possible that this is what Romer was effectively assuming in his analysis. See the final paragraph of Appendix 3.

symptomatic after a certain number of days and so self-isolate. Consider here a proportion α who self-isolate after a number of days d_0 out of the total number of days of infection d. We can approximate an adjusted value of α, which we call α'. This is the probability that someone who is infected self isolates on any particular day (independently of any test or of reaching the end of their infectious period), such that at the end of the infectious period the chance the individual has self-isolated is α, namely;

(3) $\alpha' = \alpha/d$

We then introduce α' in Equation (1) to create a new condition that the testing rate, t, must satisfy for a population following this self-isolation rule:

(4) $\{t(1-n) + [1-t(1-n)]/d + (1-t(1-n)) (1-\alpha')/d\} > R_0 / d.$

This means that the threshold testing rate is now given by

(5) $t^* = (R_0 - d\alpha' - 1 - \alpha') / [(d - d\alpha + \alpha' - 1)(1-n)]$

Suppose, as in the previous case, that $R_0 = 2.5$, $d = 14$, and $n = 0.3$. Then for a population for whom the first 5 days are asymptomatic, who then become symptomatic and self-isolate with probability $\alpha = 0.5$, Equation (4) produces vale for α'. This leads, using Equation (5) to a correspondingly reduced threshold testing rate of $t^* = 11.0$ percent. In other words, this method says that, to get R' less than 1 by randomly testing the whole population, one would only need to test 11 percent of the population because some of the population will self-isolate after 5 days.

This is a smaller value than what we found for t^* in Section 3, in this case with $\alpha = 0.5$, using the full dynamic model. The result there was that $t^* = 13$ percent. A discrepancy between these two results arises partly for the same reason that it did in the case in which $\alpha = 0$: the simple calculation carried out in both cases ignores the fact that, as time passes, testing will remove some of the infected people, so that they are no longer available to be tested on later days. But in addition, in this case here with $\alpha = 0.5$ we are making a second simplifying assumption, that there is a 'random' self-isolation process α' each day, such that at d = 14 days the chance that someone has self-isolated is exactly equal to α. This is as opposed to the detailed model in Section 3 in which, after day 5, there is a step of size α in the chance of someone self-isolating, so that nobody self-isolates for the first 5 days and then a proportion α self-isolate for a 9 day period with certainty.

Nevertheless, despite these two simplifying assumptions, it might still be useful to carry out the simple calculation described here, even despite the fact that, in this paper, we are recommending using periodic testing rather than the random testing being discussed here. This is because, as we have shown, periodic testing is more efficient than random testing. As

a result, calculating the amount of testing needed to bring R' down to 1, if testing were to be random, might well provide a *lower bound* for the threshold testing rate when using periodic testing. It might be useful to have a simple way of calculating this lower bound, even although the method of calculation is only an approximate one. We say this in the light of the current uncertainty about the true value of α in populations and the strong impact that this will have on the infection rate. In such circumstances, it seems helpful to have a calculation for $t*$ which one can carry out quickly for different values of α, without having to solve for $t*$ over and over again, using the complex non-linear approach presented in Section 3.