Oxford Review of Economic Policy, Volume 36, Number S1, 2020, pp. S77–S93

Adaptive targeted infectious disease testing

Maximilian Kasy* and Alexander Teytelboym**

Abstract: We show how to efficiently use costly testing resources in an epidemic, when testing outcomes can be used to make quarantine decisions. If the costs of false quarantine and false release exceed the cost of testing, the optimal myopic testing policy targets individuals with an intermediate likelihood of being infected. A high cost of false release means that testing is optimal for individuals with a low probability of infection, and a high cost of false quarantine means that testing is optimal for individuals with a high probability of infection. If individuals arrive over time, the policy-maker faces a dynamic trade-off: using tests for individuals for whom testing yields the maximum immediate benefit vs spreading out testing capacity across the population to learn prevalence rates thereby benefiting later individuals. We describe a simple policy that is nearly optimal from a dynamic perspective. We briefly discuss practical aspects of implementing our proposed policy, including imperfect testing technology, appropriate choice of prior, and non-stationarity of the prevalence rate.

Keywords: COVID-19, testing, adaptive learning, value of information, cost–benefit analysis
JEL classification: C9, D61, D83

I. Introduction

We have a simple message for all countries: test, test, test. Test every suspected case. If they test positive, isolate them and find out who they have been in close contact with up to 2 days before they developed symptoms, and test those people too . . . Once again, our key message is: test, test, test. (Tedros Adhanom Ghebreyesus, WHO Director-General's opening remarks at the media briefing on COVID-19, 16 March 2020).

Testing is a critical part of a response to an epidemic. At an individual level, testing allows authorities to identify and quarantine sick people, thereby stopping the spread of the disease. At a country level, testing helps authorities keep track of the disease spread, make decisions about social distancing rules, and plan for provision of supplies. However, during a sudden epidemic, such as COVID-19, testing resources can be limited (Gupta, 2020). Evidence from the COVID-19 pandemic suggests that countries

** University of Oxford; e-mail: alexander.teytelboym@economics.ox.ac.uk
*University of Oxford; e-mail: maximilian.kasy@economics.ox.ac.uk

We would like to thank Rachael Meager and Anders Kock for their careful reading of our paper. We are also grateful to Jonathan Colmer, Simon Quinn, and David Vines for their comments.
doi:10.1093/oxrep/graa018

had very different testing capacities (Hasell *et al.*, 2020). One way to think about the cost of a test is in terms of the value of testing the best possible alternative person. In other words, testing has a (shadow) cost because capacity might be difficult or impossible to ramp up quickly. Even if testing kits themselves are cheap, large-scale laboratory testing capacity might be infeasible (Hope, 2020) or it might be difficult to quickly reach all those who need testing (Weaver and Ballhaus, 2020). In this paper, we offer a simple framework that formalizes the key trade-offs that policy-makers might face under limited testing capacity, and propose an adaptive policy that can help them allocate their testing capacity as effectively as possible.

Throughout the paper, we work under the assumption that the policy-maker's objective is to minimize the total cost of disease spread. These costs are of many kinds: cost of human lives, cost of lost labour income, cost of testing kits, and reputation cost of unnecessary quarantine.

In our model, potentially sick individuals arrive over time. People might show up at the hospital because they think they have the relevant symptoms, or doctors might go out to survey and actively test people. The policy-maker can take one of three actions, for each individual:

- test the individual
  - if the individual tests positive, they are quarantined
  - if the individual tests negative, they are not quarantined;
- not test the individual, but quarantine them;
- not test the individual and release (i.e. not quarantine) them.

In our model, the policy-maker observes characteristics of individuals. These characteristics can be health-related: whether the individual has the relevant symptoms or whether the individual has been in contact with others who have symptoms or have tested positive. The characteristics can also be observables relevant to the social and economic cost of the disease: whether the individual is in a critical occupation, whether the individual has child-care responsibilities etc.

In our model, the policy-maker has access to a statistical model that can estimate the probability of the individual's having the disease conditional on the observables. The policy-maker can therefore assess the overall expected costs associated with quarantining or releasing the individual. The statistical model is imperfect in the sense that it cannot perfectly predict whether a given individual is infected based on their characteristics.

We assume that the test is perfect, but costly, so it is not possible to test everyone.[1]

If the policy-maker decides to test an individual, they incur a testing cost, but they will subsequently take an optimal quarantining decision because the test is perfect. If the policy-maker decides not to test the individual, they can make one of two costly errors:

- false quarantine: quarantining an individual who is not infected;
- false release: not quarantining an individual who is infected.

[1] This assumption does not affect the key messages of this paper as we show in section V. For a discussion of these issues, see Galeotti *et al.* (2020).

In summary, the policy-maker faces three types of costs:

– cost of testing (marginal cost or the cost of relaxing the capacity constraint);
– cost of false quarantine (e.g. forgone economic output, social isolation);
– cost of false release (e.g. spreading disease to others, not receiving early treatment).

During the COVID-19 pandemic, countries appear to have used their testing capacity in different ways. For example, there is substantial variation in the number of confirmed cases per test even after controlling for prevalence and testing capacity (Hasell *et al.*, 2020). The question we answer in this paper is: what is the testing policy that minimizes the overall costs?

The following example elucidates the key trade-offs. Suppose that the policy-maker only has 10,000 testing kits, but there are 20,000 individuals who have arrived at the hospital. Whom should the policy-maker test? Consider two policies that have been used repeatedly in the current pandemic.

Priority testing: Rank all individuals according to how likely they are to have the disease. Then test 10,000 people who are most likely to have the disease.

Several countries, such as the United States and United Kingdom, implicitly used the priority testing policy during the initial stages of COVID-19 pandemic by restricting testing to patients with strong symptoms, to those who travelled to an infected area, or to those who had been in contact with infected people (Padula, 2020).

Priority testing might be the optimal policy only if the cost of falsely quarantining individuals who are not infected is extremely high. But by testing individuals who are likely to have the disease, the policy-maker could potentially be 'wasting' tests: if the cost of a false quarantine error is not too high, the people with the highest estimated likelihood of the disease could be quarantined without testing. During the COVID-19 pandemic, many countries eventually followed this logic and advised that anyone who had symptoms or who lived with someone who had symptoms of COVID-19 must self-isolate without testing for an extended period.

Random testing: Test 10,000 individuals at random.

During the COVID-19 pandemic, a few countries and cities used random testing and there have been several calls to expand random testing (Oster, 2020; Padula, 2020). Random testing is a sensible policy if tests are very cheap. The policy-maker can learn the prevalence of the disease (thereby being able to make better decisions about testing of individuals in the future), but most people tested will not be infected. Therefore, many tests will, once again, be 'wasted'.

We proceed as follows. In section III, we point out that to make optimal decisions about testing the policy-maker needs to trade off the costs of false quarantine and false release relative to the cost of testing. Under fairly mild conditions, the optimal myopic testing policy is to test individuals with an *intermediate* likelihood of the disease. Priority testing is therefore not myopically optimal in general because the policy-maker would prefer to quarantine individuals with a high likelihood of infection without testing them and would not test or quarantine individuals who are very unlikely to be infected.

In section IV, we look at how the policy-maker's problem changes when she cares about the future. In this case, we show that the policy-maker will not initially want to

follow the myopic policy. Rather the policy-maker would want to 'explore' by initially spreading out some of her testing capacity and sacrificing some immediate benefit. The reason is that such exploratory testing gives the policy-maker valuable information about the prevalence of the disease which she can use to make better decisions about the testing of future individuals. A simple dynamic testing policy due to Thompson (1933) tells the policy-maker how much exploration is (nearly) optimal. The Thompson policy starts by initial exploratory testing. If the prevalence rate is stable, the pay-off to exploration disappears as the number of tested individuals grows because disease prevalence becomes precisely estimated. Over time, the Thompson policy converges to the optimal myopic testing policy.

In section V, we discuss some practical implementation issues, including imperfect testing. We also emphasize that our main discussion assumes that true prevalence rates across groups do not change over time. However, in epidemics, prevalence rates can change considerably. We sketch how such 'non-stationarity' can be taken into account in the context of our dynamic policies. Section VI is a conclusion.

A number of papers have considered effective testing policies. For example, in the context of HIV testing, Boozer and Philipson (2000) point out that individuals altered their transmission behaviour only if the test changed their beliefs about whether they were infected. The optimal myopic policy in our model is based on this insight about the value of information (see also Ely *et al.* (2020)), however, we also give a description of the near-optimal dynamic testing policy based on the Thompson algorithm (see, for example. Russo *et al.*, 2018; Azevedo *et al.*, 2019). A number of other recent papers have incorporated testing into models of optimal control of disease spread (e.g. Acemoglu *et al.*, 2020; Berger *et al.*, 2020; Brotherhood *et al.*, 2020; Grassly *et al.*, 2020; Piguillem and Shi; 2020). Cleevely *et al*. (2020, this issue) analysed how stratified periodic testing improves on universal random testing. Gollier and Gossner (2020) considered the efficacy of group testing.

II. Model

(i) Policy-maker's information

Let Y_i be a binary random variable denoting whether individual i (he) is infected. Let X_i be a vector of discrete characteristics that are observable and potentially predictive of Y_i. For example, X_i could include whether or not the individual has symptoms related to the disease, whether he has travelled to an infected area, whether he has been in contact with another person who has been infected, or whether he might have already had the disease and therefore built up immunity. We say individuals with characteristics $X_i = x$ are in 'group' x.

We denote by Θ_x the true prevalence of the disease among individuals in group x; that is, the share of members of group x for whom $Y_{i=1.}$ The true prevalence is unknown and the policy-maker (she) has a prior over Θ_x.

After observing the test results of individuals who were previously tested, the policy-maker can update her prior of the prevalence of the disease in each group using Bayes'

Theorem. We denote the posterior probability that an individual from group x is infected by $\hat{y}_x$. Appendix A.1 shows how the posterior probability $\hat{y}_x$ can be calculated from a simple prior used for illustration.

The policy-maker makes her decisions having potentially observed a sequence of n individuals, their characteristics, and their outcomes if they have been tested. We denote by n_x the number of individuals from group x who have been tested and by $\bar{y}_x$ the average disease prevalence among these n_x tested individuals. We assume that Θ_x does not change over time. As a result, the policy-maker cannot obtain any further information about Θ_x once $\bar{y}_x$ are known. Constant disease prevalence over time might not be a realistic assumption, but the basic trade-offs in the model will not be affected by it. We discuss the practical consequences of changing disease prevalence over time in section V.

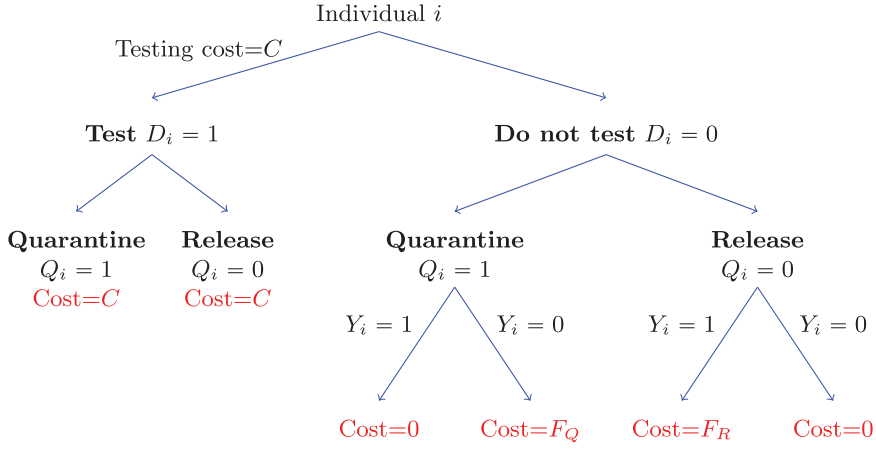
(ii) Policy-maker's choices and costs

The policy-maker's choices when she observes individual i are summarized in Figure 1.

First, the policy-maker has two choices: to test the individual ($D_i = 1$) or not to test the individual ($D_i = 0$). Testing someone for the disease comes with a cost of $C \geq 0$. Cost C can either represent the marginal cost of a testing kit or the cost of marginally relaxing the testing capacity constraint (i.e. shadow cost). The test reveals with certainty the value of Y_i and the policy-maker observes whether the individual is infected or not.[2]

Second, the policy-maker can quarantine ($Q_i = 1$) or release (not quarantine, $Q_i = 0$) the individual with or without testing. If a test has been conducted, the policy-maker can condition her quarantining decision on the observed value of Y_i.

Figure 1: Policy-maker's decisions and costs



[2] Imperfect testing does not qualitatively affect the main result (see section V).

Making wrong decisions (i.e. $Q_i \neq Y_i$) is costly. Falsely quarantining someone who is not infected comes with a cost of $F_Q > 0$.[3] Falsely releasing someone who is infected incurs a cost of $F_R > 0$. We normalize the cost of a correct decision (i.e. $Q_i = Y_i$) to 0.[4] These costs can differ by group x, but we ignore that in our notation for the sake of exposition.

III. Optimal myopic targeted testing policy

We now turn to the policy-maker's optimal myopic decision. The policy-maker takes prior beliefs as given and takes an optimal decision (D_i, Q_i) for individual i in group x having observed a sequence of (the characteristics of all) individuals, testing decisions, and the outcomes for tested individuals, i.e. $(X_j, D_j, D_j Y_j)_{j=1}^n$. This is a two-stage decision problem that can be solved by backward induction.[5]

First, consider the case where a test is conducted ($D_i = 1$) so Y_i is not observed. Recall that in this case the policy-maker can make the quarantining decision conditional on Y_i. Since $F_Q > 0$ and $F_R > 0$, the optimal decision is to set $Q_i = Y_i$. Total cost incurred in this case is the cost C of testing.

Second, consider the case where no test is conducted ($D_i = 0$) so Y_i is been observed. Recall that having observed prevalence $\bar{y}_x$ in group x, the policy-maker's posterior expectation of the prevalence in individual i's group x. is $\hat{y}_x$. Therefore, the *expected* cost of releasing an untested individual is $\hat{y}_x \cdot F_R$ while the *expected* cost of quarantining an untested individual is $(1 - \hat{y}_x) \cdot F_Q$. Hence, the optimal quarantining decision is to set $Q_i = 1$ if and only if the expected cost of a false release exceeds the expected cost of a false quarantine:

$$\hat{y}_x \cdot F_R \geq (1 - \hat{y}_x) \cdot F_Q,$$

that is, if and only if,

$$\hat{y}_x \geq \frac{F_Q}{F_Q + F_R}.$$

Note that absent a test, if we had that $F_Q = 0$ the policy-maker would quarantine everyone and if we had that $F_R = 0$ the policy-maker would not quarantine anyone. Comparing the *ex ante* expected costs of testing or not testing, we get that it is optimal to test ($D_i = 1$), if and only if

$$C \leq \min\left(\hat{y}_x \cdot F_R, (1 - \hat{y}_x) \cdot F_Q\right),$$

that is, if and only if[6]

[3] We assume that all costs are commensurate and can be measured in a single currency.

[4] The main result is not qualitatively affected by costly correct quarantine decisions (see section V).

[5] We assume that the policy-maker tests in the case when she is indifferent between testing and not testing.

[6] If the interval is empty, then no member of group x is tested.

$$\hat{y}_x \in [C/F_R, 1 - C/F_Q]\,.$$

The following proposition summarizes the policy-maker's optimal myopic policy.[7]

Proposition 1: The optimal myopic policy for individual i in group x is:

1. If $\hat{y}_x < C/F_R$, then do not test; release.
2. If $\hat{y}_x \in [C/F_R, 1 - C/F_Q]$ then test; quarantine if the test is positive; release if the test is negative.
3. If $\hat{y}_x > 1 - C/F_Q$, then do not test; quarantine.

The intuition for this result is as follows. Other things equal, if the cost F_R of a false release in a group increases, the policy-maker should start testing individuals who were previously untested (and released) because their likelihood of disease was too low.

On the other hand, if the cost of false quarantine F_Q in a group increases, the policy-maker should start testing individuals who were previously untested (and quarantined) because their likelihood of disease was too high.

Lower testing costs expand the range of disease likelihoods in which individuals are tested on both sides. If $C = 0$, all individuals are tested.

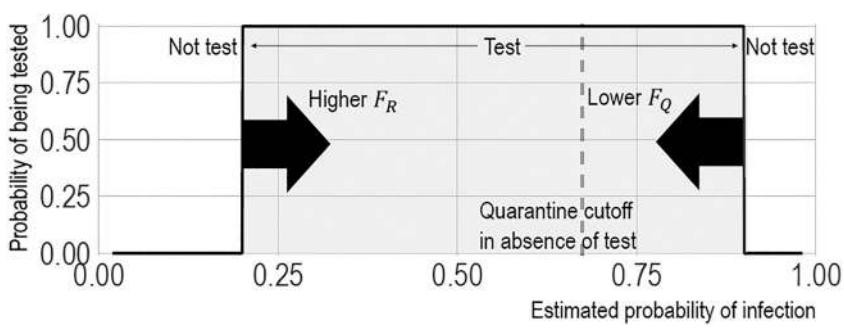
Figure 2 provides a concrete illustration of an optimal myopic testing policy. In this example, $C = 1$, $F_R = 5$ and $F_Q = 10$. If the policy-maker could not test the individual, then she would quarantine the individual if and only if the estimate of the group's disease prevalence is greater than $\frac{10}{10+5} = \frac{2}{3}$. The policy-maker tests an individual if her estimate of the group's disease prevalence is between $\frac{1}{5}$ and $\frac{9}{10}$.

IV. Adaptive targeted testing policy

The optimal myopic targeted testing policy for individual i ignores any potential value of acquired information for making future decisions. In this section, we consider what happens when the policy-maker takes this informational value into account.

Suppose that one individual arrives in every period; thus individual i arrives in period i. The policy-maker needs to make a decision (D_i, Q_i) about testing and quarantining

Figure 2: Optimal myopic testing and quarantining policy for C = 1, F_R = 5, and F_Q = 10



[7] Ties are broken in favour of testing.

this individual immediately, i.e. before the next individual arrives. When making a decision for individual i, the policy-maker has access to information $(X_j, D_j, D_j Y_j)_{j=1}^{j<i}$ for all prior individuals $j < i$.

As before when observing an individual i from group x, the policy-maker needs to make one of two decisions: (i) to test, or (ii) not to test. Denote by K_i the costs associated with the binary testing decision: (i) cost $K_i(1) = C$ of the test, and (ii) the expectation of the cost $K_i(0)$ of the error associated with either quarantining or not quarantining the individual without testing. While the cost of the test is certain, the average cost of not testing individuals from group x will depend on the prevalence Θ_x of the disease in group x; this happens because with higher Θ_x the probability of false release rises and with lower Θ_x the probability of false quarantine rises. Let $\bar{K}_x = \mathbf{E}[K_i(1) - K_i(0)|\Theta_x]$ denote the average cost difference between testing and not testing an individual in group x.

Suppose that Θ_x were known. Then the policy-maker would want to test an individual in group x if the cost of testing is lower than the expected cost of not testing, i.e. if

$$\bar{K}_x \leq 0.$$

Recall that if $\hat{y}_x < F_Q/F_R + F_Q$, then the policy-maker would release an untested individual in group x; in this case, the policy-maker would want to test if

$$\bar{K}_x = C - \mathrm{F}_R \cdot \Theta_x \leq 0.$$

Otherwise, the policy-maker would want to release an untested individual in group x; in this case, the policy-maker would test if

$$\bar{K}_x = C - F_Q \cdot (1 - \Theta_x) \leq 0.$$

If the policy-maker knew Θ_x, she would be able to take an optimal (in expectation) decision for each individual i in group x. However, the policy-maker can only form a posterior over Θ_x given the information she has access to. Armed with this posterior, the policy-maker can calculate the probability that her action is, in fact, myopically optimal (see Appendix A.2).

The policy-maker's objective is to maximize societal outcomes, i.e. to minimize cumulative expected costs over time. However, the optimal myopic testing policy that 'exploits' the full benefit of testing to the current individual (described in section III) is not dynamically optimal. The reason is that testing an individual has 'exploration' value, i.e. it allows the policy-maker to obtain a more precise estimate of the true prevalence, thereby making better decisions for individuals arriving later. As a result, the policy-maker faces an exploration–exploitation trade-off. In order to explore, the policy-maker initially spreads out some of her testing capacity and sacrifices some immediate benefit.

The optimal extent of exploration comes from a solution to a complex dynamic stochastic optimization problem. But solving for the optimal dynamic testing policy is computationally infeasible. Remarkably, however, a simple policy due to Thompson (1933) turns out to be almost optimal:

Test individual i with probability equal to the probability that $D_i = 1$ is myopically optimal.

The Thompson policy neatly captures the exploration–exploitation trade-off faced by the policy-maker. Initially, the policy-maker has little data and exploration is valuable so the probability of a testing decision for any given individual will rely heavily on the policy-maker's prior. As a result, the Thompson policy recommends to spread out testing capacity in the vicinity of the cut-offs for testing under the optimal myopic policy. As the sample size within a group x becomes large, the policy-maker learns the true prevalence Θ_x within the group and exploration becomes unnecessary. As a result, Thompson policy—as well as the dynamically optimal policy—coincides with the optimal myopic policy described in section III in the limit when n_x is large.

A spate of recent work has shown that, surprisingly, the expected total cost achieved by Thompson policy asymptotically matches the expected total costs under the fully optimal policy (Agrawal and Goyal, 2012; Kaufmann *et al.*, 2012). Therefore, in large samples the policy-maker loses almost nothing by following the Thompson algorithm (see Appendix A.3). The Thompson algorithm is used ubiquitously online in product assortment planning, revenue management, and recommendation systems by companies such as Microsoft, Google, and LinkedIn (see, for example, Russo *et al.*, 2018) and is gradually making inroads into economic policy evaluation (Kasy and Sautmann, 2019; Kasy and Teytelboym, 2020; Caria *et al.*, 2020).

(i) Illustration of myopic vs dynamic testing policies

We now illustrate how the optimal myopic testing is affected by changes in testing costs and in the costs of false quarantine/release. We also show how the Thompson policy spreads out testing compared to the optimal myopic policy.

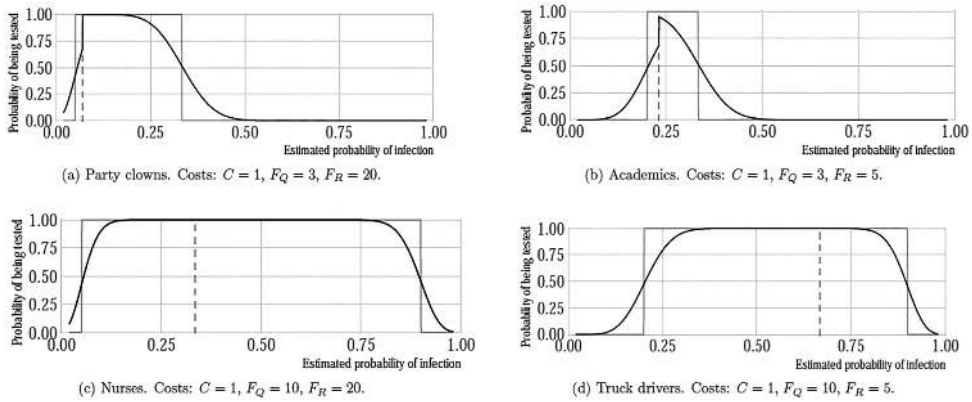
We fix $C = 1$ and $n = 50$. Let us consider four types of individuals—nurses, truck drivers, academics, party clowns—whose costs of false release and false quarantine differ as in the table below. Here we assume that the individuals' types is just one of many characteristics: the types determine the individuals' relative costs, but not their probability of infection.

	$F_R = 20$	$F_R = 5$
$F_Q = 10$	Nurse	Truck Driver
$F_Q = 3$	Party Clown	Academic

These numbers are purely illustrative; we use them for exposition and to showcase comparative statics of testing policies. Here we imagine that there are many groups of observable characteristics for every type: for example, there are nurses with and without symptoms, academics who have and who have not travelled to a conference in a disease hot spot, etc. We have already considered the optimal myopic testing policy for truck drivers in section III. The optimal myopic testing policies for all four types are illustrated in grey lines in Figures 3a–3d. For example, in an optimal myopic testing policy, the range for testing of nurses (Figure 3c) is greater than the range of testing for academics (Figure 3b) because in our illustration the costs of false quarantine and false release for nurses is greater than that of academics.

Let us first consider the Thompson testing policy for nurses and truck drivers. Figures 3c and 3d show that after 50 observations for each group the Thompson testing policy still smoothly spreads out testing around the optimal myopic testing cut-offs. In order to learn the true prevalence rate which will help make better future decisions, the

Figure 3: Optimal myopic and Thompson testing policies. Solid grey line: optimal myopic testing policy. Dashed grey line: quarantine cut-off in the absence of testing. Solid black line: Thompson policy. Number of observations (n) is 50.



(a) Party clowns. Costs: $C = 1$, $F_Q = 3$, $F_R = 20$.

(b) Academics. Costs: $C = 1$, $F_Q = 3$, $F_R = 5$.

(c) Nurses. Costs: $C = 1$, $F_Q = 10$, $F_R = 20$.

(d) Truck drivers. Costs: $C = 1$, $F_Q = 10$, $F_R = 5$.

Thompson policy suggests testing some nurses and truck drivers who would not have been tested under the optimal myopic testing policy (because their infection likelihood is either too high or too low). These tests come at the expense of lowering the probability of testing for some nurses/truck drivers who would have definitely been tested under the optimal myopic policy.

The Thompson sampling policy for party clowns and academics has two interesting further features (see Figures 3a–3b). First, the Thompson policy for academics does not recommend testing any academic with probability 1 (Figure 3b). Second, the probability of testing for academics and party clowns jumps at precisely the quarantine cut-off in the absence of a test. Intuitively, in the absence of a test the decision changes from not quarantining to quarantining around the cut-off. As a result the probability that testing is optimal also jumps.

In Appendix A.4, we illustrate the Thompson policy when the number of observations is 10 and 500. As more tests have been conducted, the policy-maker's estimate of the prevalence becomes more precise, and the rewards from exploration become smaller. As a result, the Thompson policy becomes closer and closer to the optimal myopic testing policy.

V. Extensions and implementation

Imperfect testing: Suppose that the test is imperfect with a false positive rate F_P and a false negative rate F_N. Given the estimated prevalence $\hat{y}_x$, the expected cost of testing becomes

$$C + \hat{y}_x F_N F_R + (1 - \hat{y}_x) F_P F_Q.$$

The costs of not testing remain the same. Therefore, the policy-makers test the individual if

$$C + \hat{y}_x F_N F_R + (1 - \hat{y}_x) F_P F_Q \leq \min\left(\hat{y}_x \cdot F_R, (1 - \hat{y}_x) \cdot F_Q\right),$$

that is if

$$\hat{y}_x \in \left[\frac{C + F_P F_Q}{\mathrm{F}_R - F_N F_R + F_P F_Q}, \frac{F_Q - C - F_P F_Q}{F_Q + F_N F_R - F_P F_Q}\right],$$

as long as the cut-offs are well-defined given the parameters. In general, the presence of false positive and false negative rates changes both the lower and the upper cut-offs for testing resulting in more or less testing.

Correct quarantine decision is costly: Suppose that a correct quarantine decision comes at a cost K, but the correct release decision still has a cost of zero. Then the optimal quarantining decision is to set $Q_i = 1$ if the expected cost of a false release exceeds the expected cost of a quarantine:

$$\hat{y}_x \cdot F_R \geq (1 - \hat{y}_x) \cdot F_Q + \hat{y}_x \cdot K,$$

that is, if

$$\hat{y}_x \geq \frac{F_Q}{F_Q + F_R - K}.$$

Now it is optimal to test ($D_i = 1$) if

$$C + \hat{y}_x K \leq \min\left(\hat{y}_x \cdot F_R, (1 - \hat{y}_x) \cdot F_Q + \hat{y}_x \cdot K\right),$$

that is, if

$$\hat{y}_x \in \left[{}^{C}\!\!\diagup_{F_R - K}, 1 - {}^{C}\!\!\diagup_{F_Q}\right].$$

Therefore, introducing a cost of a correct quarantine decision does not affect the upper bound for testing (as all of these individuals would have been quarantined absent a test), but it increases the lower bound for testing (as testing has become more costly).

Choice of prior: In our simulations, we have assumed a uniform prior for the prevalence rates for each group x (see Appendix A.1). In particular, the prevalence rates across groups are independent a priori. In practice, the performance of the Thompson policy will, however, depend on the choice of the prior. A practical implementation of our method will require a more sophisticated predictive model for Y based on a rich set of predictive features X, where X includes demographics, disease symptoms, contact history, etc. Such a model could for instance be constructed based on a flexible logit regression model for Y given X, with an appropriate prior for the coefficients of this regression. Alternatively, one might use a model such as those discussed in chapter 3 of Williams and Rasmussen (2006). We would assume that $\Theta_x = \frac{1}{1+e^{-g(x)}}$ and start with a Gaussian process prior $g(\cdot) \sim GP(\mu, C)$ for the function g (where μ is the mean function and C is the covariance kernel of the Gaussian process prior). For any such predictive model, we can obtain the expected posterior probability of infection $\hat{y}_x$ for individuals with characteristics x as the posterior expectation of Θ_x.

Non-stationarity: Disease prevalence rates change over time and the process is not stationary. If parameters of the model that tracks the disease spread (e.g. a SIR model) could be accurately estimated, then our methods could be adapted to learning the parameters of such a model. However, there can be a lot of disagreement among

experts about the trajectory and extent of disease spread.[8] Therefore, an alternative is to adapt our methods to an environment with non-stationary prevalence rates, which might for instance follow a rescaled random walk. The Thompson policy in such an environment would involve the model 'forgetting' older observations which might not be informative of the current state of the world (Russo *et al.*, 2018, Section 6.3). The extent of 'forgetting' would depend on the rate at which information about prevalence rates becomes obsolete (Raj and Kalyani, 2017; Besbes *et al.*, 2019).

Ethics of targeting: Targeted testing policies target. Disease prevalence as well as costs of false quarantine and false release might well vary across income, race, gender, etc. Policy-makers need to make sure that their choices of parameters and covariates do not discriminate, especially among the most vulnerable groups. These concerns are not novel to public health experts, but they can go unnoticed when decisions about individuals' lives are being made using statistical models. Our paper implies that resource allocation can be improved by carefully considering costs and benefits of testing and quarantine within any non-discrimination constraints adopted by the policy-maker.

Estimating local prevalence: How should a policy-maker maximize the precision of an estimate of the prevalence rate with a given number of tests? If the policy-maker is not concerned about the welfare of the individuals in the experimental sample, then she should sample different groups x in proportion to $\sqrt{\hat{y}_x(1-\hat{y}_x)}$, i.e. standard deviation of prevalence in the group given its prevalence rate (Neyman, 1934). Therefore, groups with prevalence closer to ½ should be sampled proportionally more. Such stratified testing strategies are particularly useful if the policy-maker subsequently uses the estimate to make a decision about a local area lockdown.

VI. Conclusion

Testing policies that use testing resources efficiently need to take into account the costs of testing, false quarantine, and false release. Our simple framework illuminates various trade-offs faced by the policy-maker when testing resources are limited. Our testing policies balance the information value of wide-ranging testing with the immediate benefit of testing and quarantining those who are likely to be infected. Practical implementation of our policies does not require any additional statistical sophistication beyond what is typically deployed to fight epidemics, but any application will require careful parameter and model calibration.

Appendix: Details of the model

A.1 Policy-maker's information

To avoid technical subtleties, we assume that all characteristics are discrete and the distribution of characteristics has finite support.

[8] Consider, for example, the difference between two highly influential models for the UK during COVID-19 due to Ferguson *et al.* (2020) and Lourencço *et al.* (2020).

Let $\Theta_x \in [0,1]$ be the prevalence of the disease among persons with characteristics ('group') $X_i = x$. The policy-maker does not precisely know the prevalence of the disease among the group. To keep things simple, let us assume that the policy-maker's prior over the prevalence is uniformly and independently distributed across groups, i.e.

$$\Theta \sim \text{Uni}\left([0,1]^k\right).$$

That is, we assume that the policy-maker essentially has no prior knowledge of the prevalence of the disease among any group and thinks that the prevalences of the disease are independent across groups. Conditional on the policy-maker's prior, the outcomes of group x follow a Bernoulli distribution with parameter Θ_x, i.e.

$$Y_i|X_i = x, \Theta \sim \text{Ber}(\Theta_x).$$

We assume that the parameters $(p_1, \ldots, p_k)$ of the distribution of characteristics are independent of the prevalence vector. Therefore, the policy-maker cannot learn about prevalence of disease simply from observing the distribution of characteristics across individuals.

The policy-maker observes a sequence $(X_j, D_j, D_j Y_j)_{j=1}^n$ of n individuals, their characteristics, and their outcomes if they have been tested. Formally, we define

$$n_x = \sum_i 1(X_i = x, D_i = 1),$$

and

$$\bar{y}_x = \frac{1}{n_x} \sum_i 1(X_i = x, D_i = 1) Y_i.$$

After observing n individuals, the policy-maker can update her prior of the prevalence of the disease in each group. Because the policy-maker's prior is uniform, the posterior has the following closed-form:

$$\Theta_x|(X_j, D_j, D_j Y_j)_{j=1}^n \sim \text{Beta}(1 + n_x \bar{y}_x, 1 + n_x(1 - \bar{y}_x)).$$

Since the mean of a random variable distributed as $\text{Beta}(\alpha, \beta)$ is $\frac{\alpha}{\alpha+\beta}$, the expected prevalence $\hat{y}_x$ in group x conditional on observing a sequence of individuals $(Y_i, X_i)_{i=1}^n$ is

$$\hat{y}_x = \mathbf{E}\left[\Theta_x|(X_j, D_j, D_j Y_j)_{j=1}^n\right] = \frac{1 + n_x \bar{y}_x}{2 + n_x}.$$

Average prevalence in the observed outcomes is sufficient to pin down the posterior prevalence, so

$$\hat{y}_x = \mathbf{E}[Y_i|X_i = x, \bar{y}_x].$$

A.2 Testing probabilities under the Thompson policy

Recall that the expected difference $\bar{K}_x = \mathbf{E}\left[K_i(1) - K_i(0) \mid \Theta_x\right]$ in costs between testing and not testing is given by

$$\bar{K}_x = C - F_R \cdot \Theta_x$$

if $\hat{y}_x < F_Q/F_R + F_Q$, and by

$$\bar{K}_x = C - F_Q \cdot (1 - \Theta_x)$$

otherwise. The posterior distribution of Θ_x is $\text{Beta}(1 + n_x\bar{y}_x, 1 + n_x(1 - \bar{y}_x))$.

Thompson sampling tests with probability equal to the posterior probability that testing is optimal. This posterior probability is given by

$$\mathbf{P}\left(\bar{K}_x \leq 0\right).$$

Let $F(y|\alpha, \beta)$ denote the cumulative distribution function of a Beta distribution with parameters α and β, evaluated at y. Then the posterior probability that testing is optimal, if $\hat{y}_x < F_Q/F_R + F_Q$, is given by

$$\mathbf{P}\left(C - F_R \cdot \Theta_x \leq 0\right) = 1 - F\left({}^{C}\!/_{F_R}; 1 + n_x\bar{y}_x, 1 + n_x(1 - \bar{y}_x)\right).$$

The posterior probability that testing is optimal, if $\hat{y}_x \geq F_Q/F_R + F_Q$, is given by

$$\mathbf{P}\left(C - F_Q \cdot (1 - \Theta_x) \leq 0\right) = 1 - F\left({}^{C}\!/_{F_Q}; 1 + n_x(1 - \bar{y}_x), 1 + n_x\bar{y}_x\right).$$

A.3 Regret bound of the Thompson algorithm

The dynamic set-up we discuss in this paper can be thought of as a so-called two-armed *contextual bandit* problem. Units arrive sequentially, we observe their group membership X_i (the 'context'), and then decide between the two 'arms' of testing ($D_i = 1$) or not ($D_i = 0$). Then rewards are realized. The rewards are $-C$ in the case of testing, and either $-F_R Y_i$ (if $\hat{y}_x < \frac{F_Q}{F_R + F_Q}$) or $-F_Q(1 - Y_i)$ (otherwise) in the case of not testing.

Conditional on the quarantining decision absent testing, this is exactly the bandit framework. We can 'concentrate out' the quarantining decision when analysing our setting. With consistency of posteriors the sign of $\hat{y}_x - \frac{F_Q}{F_R + F_Q}$ is the same as the sign of $\Theta_x - \frac{F_Q}{F_R + F_Q}$ in large samples, and theoretical results for bandit problems carry over. The learning problem for the decision-maker in this limit then becomes to figure out whether our not Θ_x lies on one or the other side of the myopic cutoffs C/F_R or $1 - C/F_Q$.

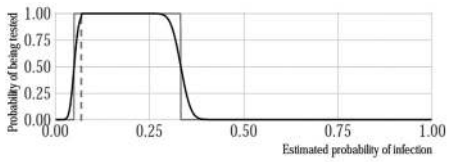
A large literature discusses guarantees for the performance of bandit algorithms (e.g. Bubeck and Cesa-Bianchi (2012) and Russo *et al.* (2018)) and more specifically of Thompson sampling (see, in particular, Agrawal and Goyal (2012) and Russo and Van Roy (2016)). These analyses bound the *regret*—the cumulative difference between realized costs and the costs that could have been achieved by taking optimal decisions given the knowledge of Θ. The guarantees in the literature are distinguished by whether they consider large sample limits for fixed Θ, or worst-case bounds over all possible Θ.

For the large sample limit case, Agrawal and Goyal (2012) show that normalized regret $T/\log(T)\bar{R}_T$ converges to a constant equal to the efficiency bound. This constant is larger the smaller the difference between treatment arms, since it then takes disproportionally longer to learn which arm is optimal. In our case, the constant is large when Θ is close to either cut-off for the myopic decision rule. We can interpret this bound to mean that only about $\log T/T$ decisions are sub-optimal, relative to those that we would have made given knowledge of Θ. For the worst-case scenario, Russo and Van Roy (2016) show that regret satisfies a finite sample prior-independent bound that grows as $\sqrt{T}$. This worst-case bound is driven by parameter values that are in a $1/\sqrt{T}$ neighbourhood of the cut-offs for the myopic rule.

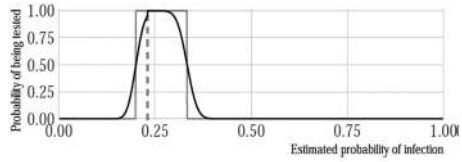
A.4 Thompson policy for $n = 500$ and $n = 10$

Figures 4a–4d illustrate the Thompson policy after 500 observations.

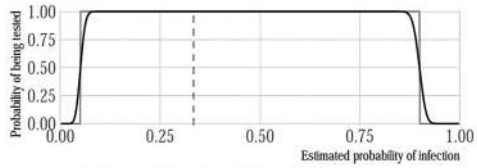
Figure 4: Optimal myopic and Thompson testing policies. Solid grey line: optimal myopic testing policy. Dashed grey line: quarantine cut-off in the absence of testing. Solid black line: Thompson policy. Number of observations (n) is 500.



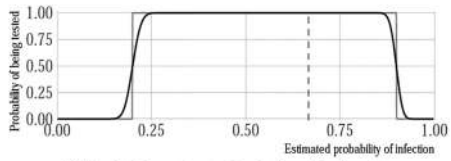
(a) Party clowns. Costs: $C = 1$, $F_Q = 3$, $F_R = 20$.



(b) Academics. Costs: $C = 1$, $F_Q = 3$, $F_R = 5$.



(c) Nurses. Costs: $C = 1$, $F_Q = 10$, $F_R = 20$.



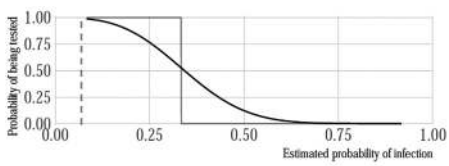
(d) Truck drivers. Costs: $C = 1$, $F_Q = 10$, $F_R = 5$.

As we showed in section A.1,

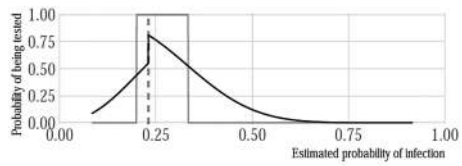
$$\hat{y}_x = \frac{1 + n_x \bar{y}_x}{2 + n_x},$$

so in small samples the support of estimated prevalence $\hat{y}_x$ can deviate significantly from the support of $\bar{y}_x$ (i.e. $[0, 1]$). Figures 5a–5d illustrate the Thompson policy after 10 observations.

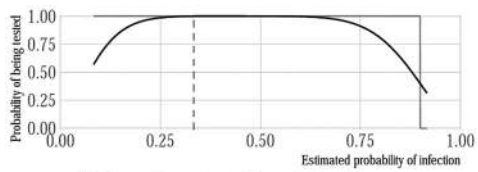
Figure 5: Optimal myopic and Thompson testing policies. Solid grey line: optimal myopic testing policy. Dashed grey line: quarantine cut-off in the absence of testing. Solid black line: Thompson policy. Number of observations (n) is 10.



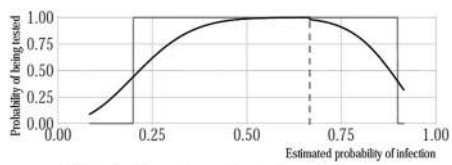
(a) Party clowns. Costs: $C = 1$, $F_Q = 3$, $F_R = 20$.



(b) Academics. Costs: $C = 1$, $F_Q = 3$, $F_R = 5$.



(c) Nurses. Costs: $C = 1$, $F_Q = 10$, $F_R = 20$.



(d) Truck drivers. Costs: $C = 1$, $F_Q = 10$, $F_R = 5$.

References

Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020), 'A Multi-risk SIR Model with Optimally Targeted Lockdown', NBER Working Paper No. 27102.

Agrawal, S., and Goyal, N. (2012), 'Analysis of Thompson Sampling for the Multi-armed Bandit Problem', in Conference on Learning Theory, **39**.1–39.26.

Azevedo, E. M., and Alex, D., Olea, J. M., Rao, M., and Weyl, E. G. (2019), 'A/B Testing with Fat Tails', available at SSRN: https://ssrn.com/abstract=3171224 or http://dx.doi.org/10.2139/ssrn.3171224

Berger, D. W., Herkenhoff, K. F., and Mongey, S. (2020), 'An SEIR Infectious Disease Model with Testing and Conditional Quarantine', NBER Working Paper No. 26901.

Besbes, O., Gur, Y., and Zeevi, A. (2019), 'Optimal Exploration–Exploitation in a Multi-armed Bandit Problem with Non-stationary Rewards', *Stochastic Systems*, 4, 319–37.

Boozer, M. A., and Philipson, T J. (2000), 'The Impact of Public Testing for Human Immunodeficiency Virus', *The Journal of Human Resources*, **35**(3), 419–46.

Brotherhood, L., Kircher, P., Santos, C., and Tertilt, M. (2020), 'An Economic Model of the Covid-19 Epidemic: The Importance of Testing and Age-specific Policies', Working Paper.

Bubeck, S., and Cesa-Bianchi, N. (2012), 'Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems', *Foundations Trends® in Machine Learning*, **5**(1), 1–122, available at http://dx.doi.org/10.1561/2200000024.

Caria, S., Gordon, G., Kasy, M., Osman, S., Quinn, S., and Teytelboym, A. (2020), 'Adaptive Treatment Assignment in Experiments for Policy Choice', mimeo.

Cleevely, M., Susskind, D., Vines, D., Vines, L., and Wills, S. (2020), 'A Workable Strategy for Covid-19 Testing: Stratified Periodic Testing rather than Universal Random Testing', *Oxford Review of Economic Policy*, **36**(Suppl), S14–S37.

Ely, J., Galeotti, A., and Steiner, J. (2020), 'Optimal Test Allocation', mimeo.

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z., Cuomo-Dannenburg, G., *et al.* (2020), 'Report 9: Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand', Imperial College COVID-19 Response Team.

Galeotti, A., Steiner, J., and Surico, P. (2020), 'An Economic Approach to Testing', mimeo, http://home.cerge-ei.cz/steiner/informationvalue.pdf

Gollier, C., and Gossner, O. (2020), 'Group Testing against Covid-19', *Covid Economics,* Centre for Economic Policy Research CEPR, **1**(2), 32–42.

Grassly, N. C., *et al.* (2020), 'Report 16—Role of Testing in COVID-19 Control', MRC Centre for Global Infectious Disease Analysis.

Gupta, P. (2020), 'Global Shortage of Coronavirus Testing Kits Critical: Experts', *Al-Jazeera*, March, https://www.aljazeera.com/news/2020/03/global-shortage-coronavirus-testing-kits-critical-experts-200323085504252.html.

Hasell, J., Ortiz-Ospina, E., Mathieu, E., Ritchie, H., Beltekian, D., Macdonald, B., and Roser, M. (2020), 'Our World In Data: Coronavirus (COVID-19) Testing', Technical Report, May, https://ourworldindata.org/coronavirus-testing.

Hope, C. (2020), '50,000 Coronavirus Tests Secretly Flown to the US after UK Lab Issues', *The Daily Telegraph*, May.

Kasy, M., and Sautmann, A. (2019), 'Adaptive Treatment Assignment in Experiments for Policy Choice', Technical Report, https://maxkasy.github.io/home/files/papers/adaptiveexperimentspolicy.pdf.

— Teytelboym, A. (2020), 'Adaptive Combinatorial Allocation Policy', mimeo.

Kaufmann, E., Korda, N., and Munos, R. (2012), 'Thompson Sampling: An Asymptotically Optimal Finite-time Analysis', in International Conference on Algorithmic Learning Theory, Springer, 199–213.

Lourenço, J., Paton, R., Ghafari, M., Kraemer, M., Thompson, C., Simmonds, P., Klenerman, P., and Gupta, S. (2020), 'Fundamental Principles of Epidemic Spread Highlight the Immediate Need for Large-scale Serological Surveys to Assess the Stage of the SARS-CoV-2 Epidemic', medRxiv.

Neyman, J. (1934), 'On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection', *Journal of the Royal Statistical Society*, **97** (4), 558–625.

Oster, E. (2020), 'We Can't Get a Handle on the Coronavirus Pandemic Without Random Testing', *Slate*, April, https://slate.com/technology/2020/04/coronavirus-research-random-testing.html.

Padula, W. V. (2020), 'Why Only Test Symptomatic Patients? Consider Random Screening for COVID-19', *Applied Health Economics and Health Policy*, April.

Piguillem, F., and Shi, L. (2020), 'Optimal COVID-19 Quarantine And Testing Policies', EIEF Working Papers Series 2004, Einaudi Institute for Economics and Finance (EIEF).

Raj, V., and Kalyani, S. (2017), 'Taming Non-stationary Bandits: A Bayesian Approach', *arXiv preprint arXiv:1707.09727* .

Russo, D. J., and Van Roy, B. (2016), 'An Information-theoretic Analysis of Thompson Sampling', *The Journal of Machine Learning Research*, **17**(1), 2442–71.

— —Kazerouni, A., Osband, I., and Wen, Z. (2018), 'A Tutorial on Thompson Sampling', *Foundations and Trends® in Machine Learning*, **11**(1), 1–96, http: //dx.doi.org/10.1561/2200000070.

Thompson, W. R. (1933), 'On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples', *Biometrika*, **25**(3/4), 285–94.

Weaver, C., and Ballhaus, R. (2020), 'Coronavirus Testing Hampered by Disarray, Shortages, Backlogs', *Wall Street Journal*, April, https://www.wsj.com/articles/coronavirus-testing-hampered-by-disarray-shortages-backlogs-11587328441.

Williams, C. K. L., and Rasmussen, C. E (2006), *Gaussian Processes for Machine Learning*, Cambridge, MA, MIT Press.